Riparian soil temperature modification of the relationship between flow and dissolved organic carbon concentration in a boreal stream

Winterdahl, M ; Futter, M ; Köhler, S ; Laudon, H ; Seibert, Jan ; Bishop, K

Abstract: Discharge is often strongly correlated to the temporal variability of dissolved organic carbon concentrations ([DOC]) in watercourses. One recently proposed way to model this is the riparian flow-concentration integration model (RIM) concept that accounts for the role of flow pathway control on [DOC] dynamics in streams. However, in boreal systems, there is also commonly a seasonal pattern, which cannot be explained by variability in discharge alone. The objectives with this study were to (1) demonstrate RIM as a tool for studying variability in stream water chemistry, (2) investigate factors related to stream water DOC variability, and (3) modify RIM to account for these factors. RIM was used with 14 years of daily discharge and almost 500 stream measurements of [DOC] from a forested boreal headwater stream. We used the calibrated RIM to account for discharge influences and then investigated variables that could be related to DOC variability (air and soil temperature, soil moisture, precipitation, antecedent flow and stream sulfate). Five alternative formulations of RIM, with temporally varying soil concentration profiles based on the variability in soil temperature and/or antecedent flow, were evaluated. The model where only the effects of riparian soil temperature on dynamics in DOC depth profiles were included performed best overall. This dynamic RIM improved the Nash-Sutcliffe to 0.58 compared to 0.42 for the flow-only formulation and reduced the median absolute error from 3.0 to 2.1 mg L−1. This study demonstrates that RIM is a simple way of modeling stream DOC and exploring controls on stream water chemistry.

DOI: https://doi.org/10.1029/2010WR010235

Riparian soil temperature modification of the relationship between flow and dissolved organic carbon concentration in a boreal stream

Mattias Winterdahl,[1] Martyn Futter,[1] Stephan Köhler,[1] Hjalmar Laudon,[2] Jan Seibert,[3,4] and Kevin Bishop[1,4]

Received 12 November 2010; revised 29 June 2011; accepted 12 July 2011; published 27 August 2011.

[1] Discharge is often strongly correlated to the temporal variability of dissolved organic carbon concentrations ([DOC]) in watercourses. One recently proposed way to model this is the riparian flow-concentration integration model (RIM) concept that accounts for the role of flow pathway control on [DOC] dynamics in streams. However, in boreal systems, there is also commonly a seasonal pattern, which cannot be explained by variability in discharge alone. The objectives with this study were to (1) demonstrate RIM as a tool for studying variability in stream water chemistry, (2) investigate factors related to stream water DOC variability, and (3) modify RIM to account for these factors. RIM was used with 14 years of daily discharge and almost 500 stream measurements of [DOC] from a forested boreal headwater stream. We used the calibrated RIM to account for discharge influences and then investigated variables that could be related to DOC variability (air and soil temperature, soil moisture, precipitation, antecedent flow and stream sulfate). Five alternative formulations of RIM, with temporally varying soil concentration profiles based on the variability in soil temperature and/or antecedent flow, were evaluated. The model where only the effects of riparian soil temperature on dynamics in DOC depth profiles were included performed best overall. This dynamic RIM improved the Nash-Sutcliffe to 0.58 compared to 0.42 for the flow-only formulation and reduced the median absolute error from 3.0 to 2.1 mg L^{-1}. This study demonstrates that RIM is a simple way of modeling stream DOC and exploring controls on stream water chemistry.

Citation: Winterdahl, M., M. Futter, S. Köhler, H. Laudon, J. Seibert, and K. Bishop (2011), Riparian soil temperature modification of the relationship between flow and dissolved organic carbon concentration in a boreal stream, *Water Resour. Res., 47*, W08532, doi:10.1029/2010WR010235.

1. Introduction

[2] Natural organic matter (NOM) influences the dynamics of many chemical species and properties, e.g., transport and toxicity of metals [*Ravichandran*, 2004; *Tipping*, 2002] and organic pollutants [*Latch and McNeill*, 2006; *Yang et al.*, 2006], acidity [*Eshleman and Hemond*, 1985; *Hruska et al.*, 2001], light penetration in lakes and streams [*Bertilsson and Tranvik*, 2000; *Karlsson et al.*, 2009], and energy mobilization in natural waters [*Jansson et al.*, 2007]. Organic carbon in the form of dissolved organic carbon (DOC, defined as the fraction of the organic carbon that passes through a filter with a pore size of about 0.2–0.7 μm) or total organic carbon (TOC, defined as the sum of DOC and particulate organic carbon (POC)) is commonly used as a proxy for NOM, which usually contains about 50%–66% carbon [*Schulten and Schnitzer*, 1993]. Several factors have been proposed as controls on surface water DOC concentrations ([DOC]), including hydrology, wetland area, climate, and atmospheric deposition [*Christ and David*, 1996; *Hope et al.*, 1994; *Kalbitz et al.*, 2000; *Monteith et al.*, 2007; *Roulet and Moore*, 2006]. Perhaps the most important mechanism for controlling DOC dynamics in many streams and rivers is discharge. A substantial part of the total annual export of organic carbon is transported during high-flow events such as storms and snowmelt [*Eimers et al.*, 2008; *Finlay et al.*, 2006; *Laudon et al.*, 2004a]. The response to discharge differs depending on dominant landscape characteristics. Many forest-dominated catchments respond to an increase in runoff with increasing [DOC], while some wetland-dominated catchments display a negative flow-[DOC] relationship [*Laudon et al.*, 2004a; *Finlay et al.*, 2006; *Eimers et al.*, 2008]. The [DOC] response to discharge has also been shown to vary across seasons, with summer and fall concentrations being higher than spring concentrations during similar discharge conditions [*Dawson et al.*, 2008; *Köhler et al.*, 2008; *Seibert et al.*, 2009]. *Ågren et al.* [2008b] hypothesized that higher [DOC] were due to higher temperatures during summer and fall.

[3] While discharge often is the first-order control on daily variability in [DOC] in forested catchments, a number of other potential drivers of the temporal variability have been suggested. Long-term variability has been hypothesized

[1]Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden.
[2]Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, Umeå, Sweden.
[3]Department of Geography, University of Zurich, Zurich, Switzerland.
[4]Department of Earth Sciences, Uppsala University, Uppsala, Sweden.

0043-1397/11/2010WR010235

to depend on mean temperature [*Freeman et al.*, 2001], atmospheric deposition of acidifying anions such as sulfate and chloride [*Evans et al.*, 2006; *Monteith et al.*, 2007] or different nitrogen compounds [*Evans et al.*, 2008; *Findlay*, 2005], and increasing atmospheric concentration of CO_2 [*Freeman et al.*, 2004]. In addition, drought-induced acidification along with variability in soil temperature, soil moisture, and extent of winter soil frost have been shown to cause short-term variability in [DOC] [*Christ and David*, 1996; *Clark et al.*, 2005; *Fröberg et al.*, 2006; *Haei et al.*, 2010; *Ågren et al.*, 2010]. A common observation is that the peak in [DOC], e.g., in spring melt, precedes the discharge peak [*Hornberger et al.*, 1994; *Laudon et al.*, 2004a], sometimes by as much as a month. This has been hypothesized to be a result of the flushing of DOC stored in the soil [*Boyer et al.*, 2000; *Hornberger et al.*, 1994; *Ågren et al.*, 2008a], indicating that antecedent flow can be an important driver for [DOC] dynamics.

[4] There are numerous models for carbon and DOC dynamics in soils [*Currie and Aber*, 1997; *Neff and Asner*, 2001; *Michalzik et al.*, 2003], but there is a rather limited number of models for simulating DOC variations in streams and watercourses [*Boyer et al.*, 2000; *Futter et al.*, 2007; *Grieve*, 1991; *Seibert et al.*, 2009; *Yurova et al.*, 2008], despite the importance of DOC in natural aquatic chemistry. With the recent interest in the observed [DOC] increase in rivers, streams, and lakes in both North America and Europe during the last several decades [*Erlandsson et al.*, 2008; *Monteith et al.*, 2007; *Worrall et al.*, 2004], models should be a way of testing different hypotheses about the causes of this increase as well as identifying the sensitivity of DOC dynamics to variation in climate and weather.

[5] Several studies have identified the riparian zone as where key processes controlling the short-term hydrogeochemical variability of runoff from forested catchments is localized [*Dosskey and Bertsch*, 1994; *Fiebig et al.*, 1990; *Hinton et al.*, 1998; *Vidon et al.*, 2010]. On the basis of empirical results from boreal forests [*Bishop*, 1991; *Bishop et al.*, 1995], *Bishop* et al. [2004] proposed a relatively simple perceptual model for how riparian soils control DOC and some other solutes in boreal headwater streams. This model is a combination of the soil solution chemical concentration profile in the riparian zone with the transmissivity feedback hypothesis of runoff generation [*Bishop*, 1991; *Kendall et al.*, 1999; *Weiler and McDonnell*, 2006]. It was proposed as a resolution to the double paradox [*Kirchner*, 2003], i.e., how catchments, on one hand, can mobilize large quantities of preevent water at storm events, while on the other hand, the chemistry of this preevent water can vary with discharge. The perceptual model introduced by *Bishop et al.* [2004] and mathematically formalized by *Seibert et al.* [2009] is a semiempirical modeling framework for simulating stream water chemistry dynamics. In the original formulation of the riparian flow-concentration integration model (RIM) the riparian soil behaves as a vertically differentiated array of chemostats [*Seibert et al.*, 2009]. These chemostats instantaneously impose a vertical concentration profile on groundwater passing laterally through a one-dimensional representation of the riparian zone (Figure 1). Superimposing the vertical profile of lateral water flow upon this "template" of the vertical soil solution concentration profile yields the chemistry of runoff water leaving the riparian zone, hillslope, or catchment at any given point in time.

[6] *Seibert et al.* [2009] used RIM "backward" by calculating soil solution concentration profiles from separate occasions of observed discharge and stream [DOC]. This identified a seasonal variability in the soil solution [DOC] profile. However, no attempt was made to parameterize this temporal variation. When using RIM to continuously model the temporal [DOC] dynamics in streams [*Köhler et al.*, 2009], it is possible to capture most of the variability driven by discharge. However, with a time-invariant soil solution concentration profile, RIM is not able to reproduce the seasonal variability in [DOC] discussed above. Several studies have shown that the [DOC] in soil solution varies during the year [*Clark et al.*, 2005; *Fröberg et al.*, 2006; *Lumsdon et al.*, 2005; *Seibert et al.*, 2009]. Possible causes of this seasonality include variable DOC production due to variations in biological activity associated with soil temperatures and/or soil moisture during different seasons, soil frost during winter, or antecedent flow effects [*Christ and David*, 1996; *Haei et al.*, 2010; *Kalbitz et al.*, 2000; *Köhler et al.*, 2009]. Seasonal changes in the concentration profile of the riparian zone are of particular importance for climate change effects.

[7] This paper has three objectives. The first is to demonstrate the potential of RIM as a "parameter-parsimonious" modeling approach for exploring chemical dynamics in streams where the riparian zone or some other interface between the terrestrial and aquatic system is important for what enters aquatic systems. The second objective is to investigate possible correlates to changes in the temporal variability in [DOC] in a forested boreal stream. Third, we propose changes, based on the potential correlates, in the structure of RIM for modeling DOC in forested boreal headwater streams that will better capture the temporal dynamics of [DOC] driven by factors in addition to flow.

2. Materials and Methods

2.1. Study Area

[8] The data used for conditioning and testing of the models in this study are from the 13 ha Västrabäcken headwater catchment, a subcatchment of the 50 ha Svartberget catchment in the Krycklan drainage basin. Both catchments are part of the Svartberget Long Term Ecological Research (64°14′N, 19°46′E) ~50 km west of the Baltic Sea coast in northern Sweden. The catchment is forested with a typical mix of species for the Swedish boreal region, with Norway spruce (*Picea abies*) in lower, moister areas and Scots pine (*Pinus sylvestris*) in upper, drier areas. Riparian areas are often covered with extensive mats of *Sphagnum* sp. mosses. The bedrock is granitic and gneissic and overlain by glacial till with thicknesses up to 10–15 m. The upslope soils are primarily well-developed iron podzols, with ~0.5 m thick organic-rich soils in the riparian areas. The elevation of the catchment ranges between 235 and 310 m, with gentle slopes inclining up to between 5% and 10%.

[9] Mean air temperature for the period 1981–2007 at the Svartberget Research Station (~1 km from the Västrabäcken catchment) is close to 2°C, and mean annual precipitation is 640 mm (period 1981–2007), of which ~35% falls as snow [*Löfvenius et al.*, 2003]. Annual discharge in Västrabäcken

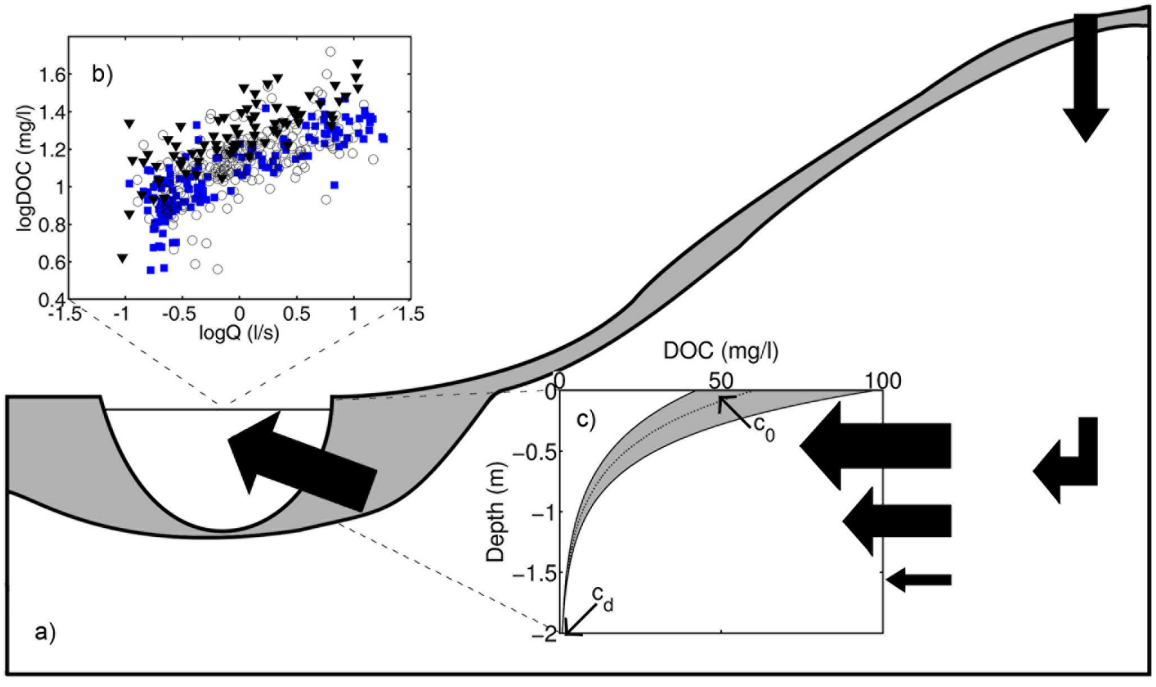

Figure 1. (a–c) Conceptual diagram of the riparian flow-concentration integration model (RIM) where the dynamics in stream dissolved organic carbon (DOC) concentration are the result of a combination of the soil solution concentration profile in the riparian soil (Figure 1b) and lateral flow across that profile at different depths in the soil (Figure 1c). Black block arrows indicate water flow and are sized relative to the volume of flow. The different symbols in Figure 1b indicate different soil temperature regimes, where blue squares indicate soil temperature $\leq$1°C, black circles indicate 1°C < soil temperature $\leq$6°C, and black triangles indicate soil temperature >6°C. The gray area in Figure 1c demonstrates the range of soil solution concentration profiles simulated by the RIM_{temp} model in this study.

is around 323 mm, and 43% of the total annual discharge is during spring flood (April and May, 1981–2007). Isotope hydrograph separations have shown that about 60%–80% of the discharge during spring flood is preevent water [*Laudon et al.*, 2002, 2007; *Rodhe*, 1987, 1989], while more than 80% of discharge in rain-driven events is preevent water [*Bishop*, 1991; *Bishop et al.*, 1990]. The stream was straightened and deepened during the 1920s to improve drainage and forest productivity, a common practice for Fenno-Scandian forest streams at the time. More detailed descriptions of the catchment have been given by *Bishop et al.* [1994] and *Köhler et al.* [2008].

2.2. Field Instrumentation and Sampling

[10] Discharge has been measured continuously since 1981. Hourly stream discharge was calculated from continuous measurements of stream water height at a 90° V notch weir, calibrated with bucket measurements and salt dilution. The water level behind the weir was measured continuously with a pressure transducer (Druck PDCR 1830), and depth data were recorded every 15 min with a Campbell Scientific data logger (model CR10x). The location of the weir is at the outlet from the Svartberget catchment and slightly downstream from the location of stream water chemistry sampling on the Västrabäcken tributary. Mean daily discharge for Västrabäcken was calculated assuming that the specific discharge was the same in the two catchments. There are differences in specific discharge between the two catchments, but earlier studies have shown that these are on the order of ±10%, mainly during high-flow events [*Bishop*, 1991]. A transect with four soil profile monitoring sites along the local groundwater flow gradient and at different distances from the stream (4, 12, 22, and 28 m) is located within the Västrabäcken catchment. Each profile is instrumented with porous ceramic suction lysimeters (P80) at six different soil depths to monitor soil water chemistry and a pressure transducer for continuous (since 2004) monitoring of the groundwater level [*Laudon et al.*, 2004b; *Nyberg et al.*, 2001]. In this study we used soil data from the S4 profile, which is in the riparian zone, 4 m from the stream. Soil water was sampled within the riparian soil at six different depths (10, 25, 35, 45, 55, and 65 cm below the ground surface). Soil temperature was measured every fourth hour with thermistors (at 5, 10, 15, 30, 50, and 70 cm below the ground surface), and soil moisture was measured every fourth hour using time domain reflectometry (TDR, Campbell Scientific, at 5, 15, 30, 40, 50, and 60 cm below the ground surface) connected to a Campbell Scientific data logger. Both soil temperature and soil moisture data used in this study were mean daily values. The highest monthly mean soil temperatures were observed in August (mean soil temperature was 8°C at 30 cm depth), and the lowest was observed in April (0.3°C at 30 cm depth). The variation in monthly mean soil moisture (TDR) was low. The highest monthly mean values were observed in December (48% at 30 cm depth), while the lowest were

observed in August (41% at 30 cm depth). The median groundwater table position during the study period was at ~46 cm below the ground surface (10th percentile was 52 cm, and 90th percentile was 41 cm).

[11] Stream water was sampled with grab samples weekly to biweekly during summer and fall, with more intense sampling during spring flood and some heavy rainfall events. The frequency went down to monthly during some winter months. Samples were frozen immediately after collection and were then analyzed on a Dohrmann carbon analyzer (1993–1994) or a Shimadzu TOC-5000 (after 1994). Both stream and soil water were usually analyzed for TOC, but in these systems, there are often no statistically significant differences between TOC and DOC, indicating that the dissolved phase is the dominant fraction [*Gadmar et al.*, 2002; *Laudon et al.*, 2004a]. Hence, we will use the term DOC in this study. For more information about field sampling and chemical analyses, see *Köhler et al.* [2008].

[12] The record of stream DOC samples stretches from 1993 to 2006 ($n = 485$, mean [DOC] = 14.9 mg L^{-1}, median [DOC] = 14 mg L^{-1}, and standard deviation is 6.7 mg L^{-1}). A linear regression between the logarithm of [DOC] and the logarithm of discharge showed that discharge explained 49% of the variance in [DOC] (Figure 1b). Discharge, precipitation, and air temperature data were available on a daily resolution for the entire study period, with the exception of short gaps caused by malfunctioning equipment. Soil temperature and soil moisture have been measured daily for more than half of the period 1995–2002.

2.3. Model Description

[13] RIM is based on the assumption that the export of chemical substances in the stream is an integrated measure of the total riparian export in the catchment. (In this study we used DOC, but a range of chemical elements or compounds is possible.) This is accomplished by superimposing lateral groundwater flow in the riparian zone on soil solution concentrations (Figure 1). By dividing the riparian soil vertically into a number of imaginary "boxes" it is possible to allow different concentrations and/or water fluxes at different levels in the soil profile. Letting the number of "boxes" approach infinity and by observing that the total export from the riparian zone is the sum of the export at every soil layer, we have

$$E = QC_{\text{stream}} = \int_z \frac{\mathrm{d}Q}{\mathrm{d}z} C_{\text{rip}}(z)\, \mathrm{d}z = \int_z q(z)\, C_{\text{rip}}(z)\, \mathrm{d}z, \qquad (1)$$

where E is the total export ($M\ T^{-1}$) from the catchment, Q is the total water flow in the stream ($L^3\ T^{-1}$), C_{stream} is the concentration in the stream ($M\ L^{-3}$), $q(z)$ is the function describing water flow at every level in the riparian soil, $C_{\text{rip}}(z)$ is the function describing soil solution chemistry profile with depth in the riparian zone, and z is depth (L) below the soil surface (depth in RIM is defined as increasing negatively downward in the soil [*Seibert et al.*, 2009]).

[14] Consistent with observations from other studies, empirical results from the transect in the Västrabäcken catchment have revealed that stream discharge is correlated with groundwater level (equation (2); Figure 2a) [*Bishop*, 1991; *Kendall et al.*, 1999; *Nyberg et al.*, 2001; *Seibert et al.*, 2003]. Data from the transect also show that [DOC] measured in the riparian soil follow a clear pattern with decreasing concentrations with depth (equation (3); Figure 2b) [*Bishop et al.*, 2004], which is found at other sites as well [*Grabs*, 2010; *Hagedorn et al.*, 2001; *Lajtha et al.*, 2005]. Regression analysis of the empirical field results showed that both flow and concentration can be described by an exponential function of depth [*Bishop*, 1991; *Seibert et al.*, 2003; *Weiler and McDonnell*, 2006]:

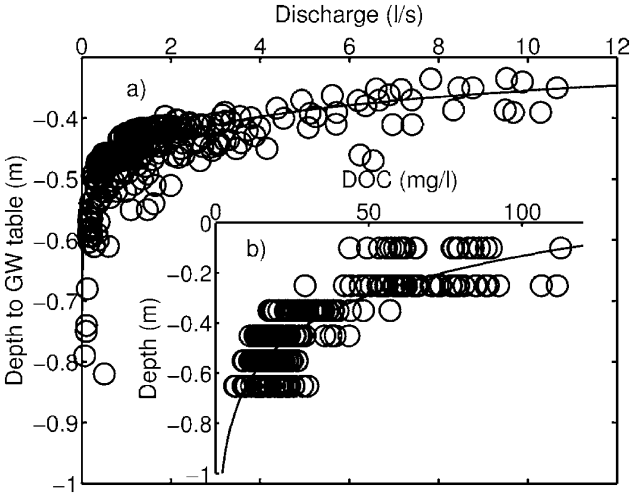

Figure 2. Observed relationships (a) between discharge in the Västrabäcken stream and depth to the groundwater table in the riparian zone 4 m from the stream ($R^2 = 0.56$, $p = 0.002$) and (b) between riparian soil solution DOC concentrations and depth in the soil ($R^2 = 0.71$, $p = 0.008$). Circles indicate observations, and the lines are best fit regressions to the data.

$$Q(z) = ke^{bz}, \qquad (2)$$

$$C_{\text{rip}}(z) = c_0 e^{fz}, \qquad (3)$$

where k and c_0 are boundary values of flow and concentration at the zero level (ground surface) and b and f are parameters describing the shape of the exponential function. (Note that the RIM framework allows the profiles to be described by other types of functions besides the exponential function.) The k and b parameters can be estimated by a regression of the natural logarithm of discharge against depth to the groundwater table in the riparian zone. The slope of that relationship is the b parameter, and the intercept is $\ln(k)$. In a similar way it is possible to estimate c_0 and f by a regression of the natural logarithm of soil solution [DOC] measured at different depths against depth in the soil. However, the preferred way to estimate the c_0 parameter is to calculate it from an estimated (fixed) base concentration c_d (Figure 1c) at an appropriately chosen base depth d [*Seibert et al.*, 2009]:

$$c_0 = c_d e^{-fd}. \qquad (4)$$

[15] By differentiating equation (2) we get

$$q(z) = \frac{\mathrm{d}}{\mathrm{d}z} Q(z) = bke^{bz} = ae^{bz}, \qquad (5)$$

where a is a new parameter defined by $a = bk$. It is now possible to rewrite equation (1):

$$E = \int_{z} q(z)\, C_{\mathrm{rip}}(z)\, \mathrm{d}z = \int_{z_0}^{z_1} a e^{bz} c_0\, e^{fz} \mathrm{d}z, \tag{6}$$

where z_1 and z_0 are the upper (the groundwater table) and lower limits, respectively. The groundwater table depth at every time step was calculated under the assumption that all runoff in the stream is routed through the riparian soil:

$$z_1 = b^{-1} \ln\left(\frac{bQ}{a}\right). \tag{7}$$

[16] This representation of flow neglects overland flow or other mechanisms where water is bypassing the riparian zone, direct precipitation input in the stream, and any other water input between the riparian zone and the stream. In this study we have used RIM in a lumped approach, assuming that one riparian soil solution profile represents the entire catchment, to focus on seasonality.

[17] Equation (6) is solved analytically [*Seibert et al.*, 2009] with discharge in the stream as input data:

$$C_{\mathrm{stream}} = \frac{c_0 (a/b)^{1-\eta}}{\eta} Q^{\eta - 1}, \tag{8}$$

where $\eta = (b + f)/b$.

2.4. Correlation Analysis

[18] The first applications of RIM for modeling [DOC] in Västrabäcken revealed that there was a need for more dynamics in the model, thereby addressing the temporal variability in the flow-[DOC] relationship [*Köhler et al.*, 2009; *Seibert et al.*, 2009]. We used a residual analysis of the calibrated original RIM (hereafter called $\mathrm{RIM_{static}}$) to investigate potential correlates to the temporal variability in [DOC]. The purpose of using $\mathrm{RIM_{static}}$ was to remove any discharge effects on the [DOC] variability, thereby allowing identification of other potential correlates. On the basis of the residual analysis we decided on possible candidate functions for a dynamic RIM (equation (9)). A conceptually attractive way of modifying RIM for modeling [DOC] would be to allow the shape of the soil solution concentration profile to vary in time; that is, the f parameter would vary in time:

$$C_{\mathrm{stream}} = Q^{-1} \int_{z_0}^{z_1} a e^{bz} c_0 e^{f(t)z} \mathrm{d}z. \tag{9}$$

[19] One could also conceive a structure where the c_d parameter as well is dependent on time ($c_d(t)$), i.e., a shift of the entire profile (or depth-averaged concentration) to higher or lower concentrations. Both approaches were investigated in this study, but with consideration given to keeping the number of extra parameters to a minimum.

[20] The residuals of $\mathrm{RIM_{static}}$ were defined as the difference between the observed concentrations O_i in the stream and the concentrations predicted by $\mathrm{RIM_{static}}$ M_i at every time step i:

$$\mathrm{res}_{\mathrm{RIM_{static}}}(i) = O_i - M_i. \tag{10}$$

[21] Residual analysis was conducted by correlating model residuals with different possible variables thought to cause temporal variability in [DOC]. In this study we used both the Pearson's correlation r and the Spearman's rank correlation ρ. Potential controls on residual variability were analyzed by attempting to correlate unexplained variability in DOC to discharge, antecedent flow (the cumulative flow, AQ, during the preceding n number of days, i.e., $\mathrm{AQ} = \sum_{i=1}^{n} Q_i$, where Q_i is specific discharge at time i and we tried n from 1 to 365 days), sulfate concentration in the stream, air temperature, precipitation, and observed riparian soil temperature and soil moisture at six different depths (soil temperature and moisture data are from the riparian S4 plot). The variables included in the analysis were chosen on the basis of their potential effect on DOC dynamics. One could imagine other potential variables to include, but the list of variables was limited by the amount of available data.

[22] With the residual analysis approach we assumed that RIM was a "correct" representation of the system and that the unexplained variability in concentrations could be modeled with a temporally variable f and/or c_d. To investigate the possibility of modeling the [DOC] variability in Västrabäcken with RIM, we back calculated the f parameters of $\mathrm{RIM_{static}}$ from observed stream concentrations at every sampling occasion [*Seibert et al.*, 2009]. It is not possible to analytically solve for f in equation (8), so a numerical approach was adopted using a combination of the bisection, secant, and quadratic interpolation methods [*Press et al.*, 2007]. The time series of these back-calculated f parameters (f_{back}) was then correlated with the residuals of $\mathrm{RIM_{static}}$.

[23] Further, we tried a forward selection stepwise regression with model residuals as the response variable and discharge, antecedent flow, sulfate concentration in the stream, air temperature, precipitation, and observed soil temperature and soil moisture as explanatory variables. This was done as an indication of possible variables to include in the dynamic version of RIM.

2.5. Dynamic RIM Candidates

[24] On the basis of the correlation analysis, new possible model structures for a modified, dynamic version of RIM were proposed (equation (9)). All the new model structures where f and/or c_d were allowed to vary included multiple linear regressions according to the general formula

$$\alpha(t) = \beta_0 + \sum_{i=1}^{m} \beta_i x_i(t), \tag{11}$$

where $\alpha(t)$ is the model describing the temporal variability in f or c_d, x_i are possible predictor variables from the correlation analysis, β_i are parameters, and m is the number of predictor variables included in the model. The function $\alpha(t)$ (equation (11)) was then incorporated into the RIM structure instead of f and/or c_d. The performance of the dynamic RIM candidates was compared with $\mathrm{RIM_{static}}$.

2.6. Soil Temperature Modeling

[25] Since the measured soil temperature record from Västrabäcken contains gaps and does not span the entire study period, soil temperature was modeled (Nash-Sutcliffe efficiency was 0.90 between measured and modeled soil temperature at 30 cm depth; median absolute error was 0.5°C) for the entire period 1993–2006 using the soil temperature model by *Rankinen et al.* [2004]. The modeled soil temperature was not used in the correlation analysis but only in model simulations when applying the different dynamic RIM models.

2.7. Parameter and Uncertainty Estimation

[26] The parameters in RIM could be estimated from observed data in the riparian zone and in the stream (Figure 2) [*Seibert et al.*, 2009]. However, when modeling an entire catchment, the effective parameters, i.e., the parameters that are valid at the modeling scale (in this case catchment scale), are needed. In order to optimize the performance of the model and to estimate the combined modeling uncertainty we therefore used the generalized likelihood uncertainty estimation (GLUE) [*Beven*, 2008; *Beven and Binley*, 1992]. We used observed stream discharge and, for some of the dynamic RIM models, modeled soil temperature as input data and observed [DOC] data in the Västrabäcken stream as conditioning data.

[27] GLUE is based on Monte Carlo simulations where a large number of parameter values are sampled randomly from predefined distributions. Since the distributions of the parameters and model residuals were unknown, we chose continuous uniform distributions for all parameters within the predefined parameter limits shown in Tables 1 and 2. The parameter limits were decided from what seemed to be reasonable limits judged from the physical meaning of the parameters (when there were any) and from preliminary tests. For example, c_d, which is the base concentration at a predefined depth d in the soil, cannot be negative, and values above 15 mg L^{-1} are unlikely. The model was not sensitive to the depth of the lower boundary d, and it was therefore set to 2 m.

[28] We used two approaches in the conditioning of the models. The first was to use an automatic GLUE estimation of all parameters, i.e., both parameters controlling the flow profile (a and b) and parameters controlling the [DOC] soil solution profile (e.g., f and c_d). The second approach was to fix the flow parameters a and b to empirically estimated values and then to automatically condition the parameters controlling the soil solution [DOC] profile. In the second approach we estimated the a (1163 L $s^{-1}m^{-1}$) and b (8.8 m^{-1}) parameters by applying a nonlinear regression between discharge in the Västrabäcken stream and groundwater table position in the riparian soil (Figure 2a) [*Bishop*, 1991; *Nyberg et al.*, 2001; *Seibert et al.*, 2003, 2009]. When the a and b parameters were included in the automatic GLUE conditioning, we used a "soft calibration" [*Seibert and McDonnell*, 2002] to constrain the mean groundwater table position to reasonable depths on the basis of observations [*Nyberg et al.*, 2001; *Seibert et al.*, 2003].

Table 1. Model Parameters, Parameter Units, Parameter Descriptions, and Parameter Ranges Used in Model Conditioning of the RIM_{static} Model[a]

Parameter	Unit	Description	Parameter Range
a	L s^{-1} m^{-1}	Change in flow at soil surface	320–45,000
b	m^{-1}	Slope of Q-GW[b] relationship	4–15
c_d	mg L^{-1}	[DOC] at base level	0–6
f	m^{-1}	Slope of soil solution [DOC] profile	0.5–4
d	m	Base level	−2

[a]RIM, riparian flow-concentration integration model; [DOC], dissolved organic carbon concentration.
[b]Q, stream discharge; GW, groundwater table position.

Table 2. Parameter Ranges Used in Model Conditioning for the Parameters in the Dynamic Versions of RIM[a]

Model	c_d	β_0	β_1	β_2	γ_0	γ_1
RIM_{temp}	0–5	0.5–4	−0.05–0.12	X	X	X
RIM_{ant}	0–5	0–4	−0.001–0.0025	X	X	X
RIM_f	0–5	0–4	−0.05–0.15	−0.002–0.002	X	X
RIM_{cf1}	X	0–3	−0.05–0.1	X	0–6	0–0.4
RIM_{cf2}	X	0–3.5	−0.002–0.002	X	0–6	−0.1–0.3

[a]Parameters not used in the model are indicated by a cross.

[29] One common practice within the GLUE methodology is to use informal likelihood measures to estimate model performance [*Beven*, 2008; *Beven and Binley*, 1992]. It has been recognized that different likelihood (or performance) measures focus on different aspects of the model performance and hence indicate different behavioral parameter sets [*Choi and Beven*, 2007; *Gupta et al.*, 1998]. Therefore, we used four different informal likelihood measures in this study: the maximum absolute error (MaxAE; equation (12)), the bias (BIAS; equation (13)), the adjusted Nash-Sutcliffe efficiency index (NS_{adj}; equation (15)), and the adjusted E_1 efficiency index ($E_{1,adj}$; equation (17)). The maximum absolute error was calculated according to

$$\mathrm{MaxAE} = \max\{\,|O_i - M_i|,\ 1 \le i \le n\,\}, \tag{12}$$

where n is the number of observations, O_i is the observed concentration at time i, and M_i is the modeled concentration at time i. The BIAS, which indicates if the model has a general tendency to overpredict or underpredict, was calculated according to

$$\mathrm{BIAS} = \frac{1}{n}\sum_{i=1}^{n}(O_i - M_i). \tag{13}$$

[30] The Nash-Sutcliffe efficiency index (NS) compares the squared errors between observed data and model results with the squared errors between observed data and the whole period average of observations at every time step i [*Nash and Sutcliffe*, 1970]:

$$\mathrm{NS} = \left(1 - \frac{\sum_{i=1}^{n}(O_i - M_i)^2}{\sum_{i=1}^{n}(O_i - \bar{O})^2}\right), \tag{14}$$

where $\bar{O}$ is the mean of the observed data. The NS varies between $-\infty$ and 1, where a model with a perfect fit to the observed data would have a NS equal to 1. A model with a negative NS performs worse than the observed mean, while a model with a NS equal to 0 performs just as good as the

observed mean. *Clarke* [2008b] proposed an adjusted version of the NS for use in model intercomparisons. The rationale for using the adjusted version would be to account for the different number of parameters used in different models:

$$\mathrm{NS}_{\mathrm{adj}} = \left(1 - \frac{\frac{1}{n-p}\sum_{i=1}^{n}(O_i - M_i)^2}{\frac{1}{n-1}\sum_{i=1}^{n}(O_i - \bar{O})^2} \right), \tag{15}$$

where p is the number of parameters in the model under study. The NS has been criticized as being sensitive to outliers [*Legates and McCabe*, 1999], and *Legates and McCabe* [1999] proposed a modified efficiency index, which they called the E_1 index:

$$E_1 = \left(1 - \frac{\sum_{i=1}^{n}|O_i - M_i|}{\sum_{i=1}^{n}|O_i - \bar{O}|} \right). \tag{16}$$

[31] We have modified the E_1 index according to the method used by *Clarke* [2008b] for the adjustment of the NS to account for the number of parameters used in different models:

$$E_{1,\mathrm{adj}} = \left(1 - \frac{\frac{1}{n-p}\sum_{i=1}^{n}|O_i - M_i|}{\frac{1}{n-1}\sum_{i=1}^{n}|O_i - \bar{O}|} \right). \tag{17}$$

[32] The different likelihood measures were sequentially combined with repeated use of Bayesian updating. Bayes' equation can be stated in the form

$$P[M(\Theta_i)|\,O_T] = \frac{P[M(\Theta_i)]\;P[O_T|\,M(\Theta_i)]}{\sum_{i=1}^{m} P[M(\Theta_i)]\;P[O_T|\,M(\Theta_i)]}, \tag{18}$$

where $P[M(\Theta_i)|O_T]$ is the posterior likelihood for model M with the ith parameter set Θ_i, conditioned on the observed data O during time period T, $P[M(\Theta_i)]$ is the prior likelihood, $P[O_T|M(\Theta_i)]$ is the likelihood (the performance of any of the likelihood measures) of simulating O_T given model M with the ith parameter set Θ_i, and m is the number of parameter sets. The denominator in equation (18) is a normalization factor that ensures that the cumulative posterior likelihood is unity. Bayesian updating will adjust the posterior likelihood of the different parameter sets according to the performance of the parameter sets measured with the four likelihood measures. The advantage of using a multiplicative combination of likelihood measures is that the posterior likelihood will be zero if any of the four likelihood measures is zero.

[33] The BIAS and MaxAE likelihoods were slightly modified before the Bayesian updating. We used the absolute value of the BIAS likelihoods, and all parameter sets with values larger than 2 were considered to be nonbehavioral in the continued analysis. The MaxAE was modified by taking its inverse. In addition, all parameter sets with $\mathrm{NS}_{\mathrm{adj}}$ or $E_{1,\mathrm{adj}}$ less than 0 were considered to be nonbehavioral. All likelihood measures were normalized to the total range of the population before the Bayesian updating to ensure that the cumulative likelihood was unity.

[34] One common practice in GLUE is to set a behavioral acceptance threshold to judge which parameter sets should be considered behavioral. Instead of using a threshold, we selected the 200 top-performing parameter sets on the basis of the four likelihood measures combined with the Bayesian updating in evaluation of each model. The population of behavioral parameter sets would then be of the same size for all considered models while still retaining an estimation of the combined modeling uncertainty. The model sensitivity to the different parameters was investigated with the Hornberger-Spear-Young global sensitivity analysis (GSA) [*Spear and Hornberger*, 1980].

[35] For the residual analysis, $\mathrm{RIM}_{\mathrm{static}}$ was conditioned on the basis of the entire data record (1993–2006). The model with the best fit to observed [DOC] on the basis of the values of the likelihood measures employed was used in the correlation analysis. In addition, we studied what could be called the GLUE residuals, i.e., the occasions when the observed concentrations were outside the uncertainty limits of the 200 top-performing models. For the investigation of the dynamic versions of RIM and the comparison with $\mathrm{RIM}_{\mathrm{static}}$, a split sample calibration scheme was employed [*Klemes*, 1986]. We used the [DOC] data from 2000–2006 in conditioning and data from 1993–1999 for testing the models.

[36] Comparing different models requires caution [*Clarke*, 2008a]. To test for differences in the performance of the 200 top-performing parameter sets for the different models, we used the Kruskal-Wallis one-way analysis of variance by ranks [*Kruskal and Wallis*, 1952]. Pairwise comparisons were done with Scheffe's method [*Scheffe*, 1953]. In order to investigate the absolute and relative errors in model-predicted concentrations we calculated the median absolute error (MedAE) and the median absolute percent error (PE) in DOC concentrations:

$$\mathrm{MedAE} = \mathrm{median}\,\{\,|O_i - M_i|,\ 1 \le i \le n\,\}, \tag{19}$$

$$\mathrm{PE} = \mathrm{median}\left\{ 100\left|\frac{O_i - M_i}{O_i}\right|,\ 1 \le i \le n \right\}. \tag{20}$$

[37] We also controlled possible parameter correlations in the different models, both visually and with the Spearman's rank correlation.

[38] To test that the models simulated [DOC] for the "right reasons" and not just because of a suitable number of parameters, we compared simulated soil solution concentrations with measured concentrations in S4. We calculated the mean absolute error between simulated and observed soil solution concentrations at the six different depths where observations were available. However, we expected that the simulated concentrations would deviate from the observed concentrations in absolute numbers since the model simulates flow weighted average riparian soil solution concentrations for the entire catchment, while observed concentrations are from a single profile. Consequently, we also wanted to compare the relative dynamics of observed and simulated concentrations alone. The standard scores (Student's t statistic)

of simulated concentrations were compared to the standard scores of observed concentrations:

$$t_i = \frac{x_i - \bar{x}}{s}, \quad (21)$$

where t_i is the standard score of simulated or observed [DOC] at time i, x_i is observed or simulated [DOC] at time i, $\bar{x}$ is the sample mean of observed or simulated [DOC], and s is the sample standard deviation of observed or simulated [DOC]. Though this procedure will corrupt differences in mean and variance of observed and simulated concentrations, it will retain the dynamics of the time series.

3. Results

3.1. Conditioning of RIM_{STATIC}

[39] RIM_{static} predicted the general [DOC] variability for 1993–2006, but the model had a tendency to underestimate during summer and fall, while it generally overestimated during spring flood. Hence, it did not simulate the observed seasonality in [DOC] variability. The BIAS of the 200 behavioral parameter sets was close to zero (−0.21 to 0.26 mg L^{-1}), while the mean $E_{1,adj}$ was 0.30 (standard deviation of 0.01) and the mean NS_{adj} was 0.44 (standard deviation of 0.01). The distribution of NS_{adj} was strongly skewed toward higher values. The seasonal pattern was also visible when investigating the residuals of the model with the highest performance (Figure 3b). The GLUE residuals revealed the same pattern, but they differed slightly in absolute numbers. The results did not improve when the a and b parameters, controlling the simulated lateral flow paths through the riparian soil, were included in model conditioning.

[40] The performance within specific years was variable. The $E_{1,adj}$ and NS_{adj} of individual years did not reveal any particular pattern over time or any correlation with, e.g., weather data. There was, however, a relationship between the BIAS ($R^2 = 0.45$) and the MaxAE ($R^2 = 0.31$) with total annual runoff. High runoff tended to result in large positive bias and large maximum errors, while low runoff resulted in smaller maximum errors and slightly negative bias. All the behavioral parameter sets displayed a clear, albeit nonlinear, correlation between the f and c_d parameters ($\rho = 0.99, p < 0.001$).

3.2. Correlation Analysis

[41] In order to find possible drivers for the temporal [DOC] variability we correlated RIM_{static} residuals with a

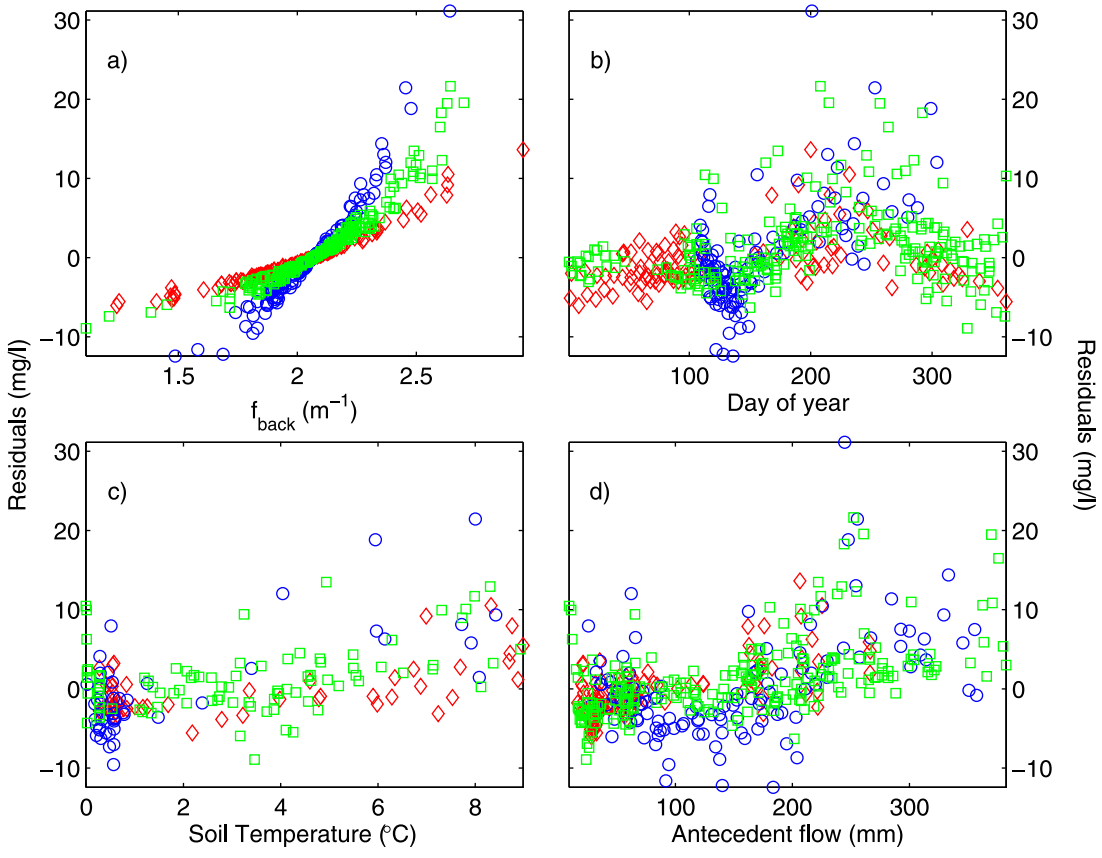

Figure 3. Scatterplots of RIM_{static} residuals against (a) back-calculated f parameters, (b) day of year, (c) observed soil temperature at 30 cm below ground surface, and (d) antecedent flow (total flow in the preceding 159 days). Red diamonds indicate a low-flow situation (low flow is defined as <25th percentile), blue circles indicate high flow (high flow is defined as >75th percentile), and green squares indicate medium flow (between the 25th and 75th percentiles).

range of potential predictors. The highest correlations were obtained for soil temperature and antecedent flow accumulated over 159–168 days (Table 3 and Figures 3c and 3d). Most of the explanatory variables correlated significantly using both correlation methods. When applying RIM_{static}, the purpose was to filter out the discharge effect on [DOC] variability, and indeed, discharge did not correlate significantly with model residuals. The correlation between model residuals and soil moisture was only significant at 5 cm ($p < 0.05$ for Spearman's correlation), 30 cm ($p < 0.05$ for both Pearson's and Spearman's correlation), and 40 cm ($p < 0.01$ for Spearman's correlation). However, the distributions for soil moisture were nearly bimodal. In addition, soil moisture (when the correlation was significant) and sulfate were the only variables to correlate negatively with the residuals. Soil temperature correlated significantly at all depths, but the highest correlations were obtained at different depths for different correlation methods. The highest Pearson's correlation was achieved at 15 cm, while the highest Spearman's ρ was obtained at 70 cm. The correlation between antecedent flow and model residuals were similar for the two methods. The total flow during 159 preceding days displayed the highest r, while the highest ρ was obtained for the total flow during 168 preceding days. Antecedent flow for the preceding 7 or 14 days correlated weakly with the residuals. Sulfate has been hypothesized to affect both long-term and short-term DOC concentrations [*Clark et al.*, 2005; *Evans et al.*, 2006; *Monteith et al.*, 2007], and we found sulfate to correlate significantly with the model residuals. However, the correlation was weaker than for soil temperature or antecedent flow and seemed to be controlled by a few outliers (not shown).

[42] The stepwise regression indicated a model where only soil temperature, soil moisture, and antecedent flow, in that order, could be considered to be significant predictors. However there was considerable covariation between some of the predictor variables (Table 4). Antecedent flow (159 days) correlated with both soil moisture ($r = 0.31$–0.66, $p < 0.001$) and soil temperature ($r = 0.69$–0.74, $p < 0.001$). Soil temperature (at all levels) only correlated with soil moisture at 5 cm.

Table 3. Correlation Statistics Between RIM_{static} Model Residuals and Predictor Variables[a]

Variable	Pearson	Spearman
f_{back}	0.88	0.97
Air temperature	0.27	0.32
Precipitation	0.11	0.09
Discharge	−0.09	0.02
Antecedent flow[b]	0.51	0.53
Soil temperature[c]	0.54	0.46
Soil moisture[d]	−0.15	−0.19
Sulfate concentration	−0.33	−0.34

[a]Model residuals are the differences between observed and RIM_{static}-predicted [DOC]. All values have significant correlation ($p < 0.05$) except those for discharge.

[b]The number is the highest correlation, which for the Pearson correlation was for the accumulated flow over 159 days. The highest Spearman correlation was for the accumulated flow over 168 days.

[c]Numbers indicate the highest correlation, which was at 15 cm for the Pearson correlation and at 70 cm for the Spearman correlation.

[d]Only significant at 5, 30, and 40 cm. Numbers indicate the highest correlations, which was at 30 cm for the Pearson correlation and at 40 cm for the Spearman correlation.

[43] The correlation between model residuals and f_{back} was high ($r = 0.88$, $\rho = 0.97$, $p < 0.001$), and there was a clear linear relationship between the residuals and f_{back} (Figure 3a). The relationship was, however, heteroscedastic; that is, the variance was not constant across the range of values. This heteroscedasticity could explain why the Spearman's ρ revealed a higher correlation than the Pearson's correlation method.

3.3. Dynamic RIM Candidates

[44] On the basis of the correlation analysis, only soil temperature and antecedent flow could be considered to be important variables. We proposed five dynamic RIM structures based on these. The different models, which could be considered to be different hypotheses about the functioning of the system, are outlined as follows:

[45] 1. The RIM_{temp} model simulates a change in the shape of the soil solution profile on the basis of variability in soil temperature alone. Hence, the hypothesis is that soil temperature is the sole driver of differentiated concentration changes in the riparian soil solution, which is reflected as DOC variability in the stream. This is modeled as f being dependent on soil temperature at 30 cm depth; that is, $f(t)$ in equation (9) takes the form $f(t) = \beta_{1,\mathrm{temp}} T_{\mathrm{soil}}(t) + \beta_{0,\mathrm{temp}}$.

[46] 2. The RIM_{ant} model is similar to RIM_{temp}, but the variability is solely dependent on antecedent flow instead of soil temperature. Hence, f is modeled as being dependent on antecedent flow during the preceding 159 days (chosen because it was the lag that gave the highest correlation to the residuals); that is, $f(t)$ takes the form $f(t) = \beta_{1,\mathrm{ant}} \sum_{i=t}^{t-159} Q_i + \beta_{0,\mathrm{ant}}$.

[47] 3. As in previous two models, RIM_f simulates variability only by changing the shape of the soil solution profile. The variability is driven by changes in both soil temperature and antecedent flow. This is modeled as f being dependent on antecedent flow during the preceding 159 days and soil temperature at 30 cm depth; that is, $f(t)$ takes the form $f(t) = \beta_{2,f} \sum_{i=t}^{t-159} Q_i + \beta_{1,f} T_{\mathrm{soil}}(t) + \beta_{0,f}$.

[48] 4. In the RIM, short-term variability can be simulated by varying only the shape of the soil solution profile (the f parameter) as in the first three models. It is, however, also plausible that the base concentration (the c_d parameter) could vary in time, resulting in a change in depth-averaged [DOC]. RIM_{cf1} simulates temporal variability by varying both the shape of the soil solution profile and the base concentration. The hypothesis is that the change in profile shape is driven solely by soil temperature and that the change in depth-averaged [DOC] is driven by antecedent flow. This is simulated as f being dependent on soil temperature at 30 cm depth and c_d being dependent on antecedent flow during the preceding 159 days; that is, $f(t)$ takes the form $f(t) = \beta_{1,\mathrm{cf1}} T_{\mathrm{soil}}(t) + \beta_{0,\mathrm{cf1}}$, and $c_d(t)$ takes the form $c_d(t) = \gamma_{1,\mathrm{cf1}} \sum_{i=t}^{t-159} Q_i + \gamma_{0,\mathrm{cf1}}$.

[49] 5. RIM_{cf2} is similar to RIM_{cf1}, but the shape of the soil solution profile varies as a function of antecedent flow, and the base concentration varies as a function of soil temperature. This is modeled as f being dependent on antecedent flow during the preceding 159 days and c_d being

Table 4. Correlation Matrix for the Predictor Variables Used in the Correlation Analysis[a]

	Q	AQ (159)	AQ (7)	AQ (30)	SO_4^{2-}	Air Temperature	Precipitation
Q	**1**	**0.33**	**0.85**	**0.46**	**−0.43**	**0.46**	−0.01
AQ (159)		**1**	**0.42**	**0.57**	**−0.79**	**0.55**	0.26
AQ (7)			**1**	**0.62**	**−0.45**	**0.48**	−0.15
AQ (30)				**1**	**−0.51**	**0.54**	0.04
SO_4^{2-}					**1**	**−0.45**	−0.16
Air temperature						**1**	0.09
Precipitation							**1**

	Soil Temperature					
	5 cm	10 cm	15 cm	30 cm	50 cm	70 cm
Q	0.21	0.21	0.21	0.22	0.22	0.24
AQ (159)	**0.74**	**0.73**	**0.72**	**0.70**	**0.71**	**0.69**
AQ (7)	0.25	0.25	0.25	0.26	0.26	0.26
AQ (30)	0.26	0.24	0.23	0.2	0.19	0.16
SO_4^{2-}	**−0.38**	**−0.36**	**−0.34**	**−0.33**	**−0.35**	**−0.34**
Air temperature	**0.50**	**0.47**	**0.45**	**0.42**	**0.41**	**0.36**
Precipitation	**0.30**	**0.30**	**0.31**	**0.31**	**0.32**	**0.33**
Soil temperature (5 cm)	**1**	**1**	**0.99**	**0.98**	**0.98**	**0.95**
Soil temperature (10 cm)		**1**	**1**	**0.99**	**0.99**	**0.97**
Soil temperature (15 cm)			**1**	**1**	**0.99**	**0.97**
Soil temperature (30 cm)				**1**	**1**	**0.99**
Soil temperature (50 cm)					**1**	**0.99**
Soil temperature (70 cm)						**1**

	Soil Moisture					
	5 cm	15 cm	30 cm	40 cm	50 cm	60 cm
Q	**0.60**	**0.45**	**0.53**	**0.68**	**0.38**	0.19
AQ (159)	**0.66**	**0.49**	**0.62**	**0.49**	**0.59**	**0.31**
AQ (7)	**0.60**	**0.42**	**0.53**	**0.66**	**0.38**	0.16
AQ (30)	**0.72**	**0.33**	**0.56**	**0.60**	**0.46**	0.13
SO_4^{2-}	**−0.48**	**−0.75**	**−0.84**	**−0.75**	**−0.87**	**−0.65**
Air temperature	**0.62**	0.23	**0.46**	**0.49**	**0.51**	0.09
Precipitation	0.16	−0.11	−0.06	−0.03	0.06	−0.16
Soil temperature (5 cm)	**0.63**	0.13	0.2	0.14	0.23	−0.05
Soil temperature (10 cm)	**0.62**	0.12	0.18	0.13	0.19	−0.07
Soil temperature (15 cm)	**0.61**	0.11	0.16	0.13	0.17	−0.07
Soil temperature (30 cm)	**0.59**	0.11	0.15	0.13	0.14	−0.07
Soil temperature (50 cm)	**0.58**	0.15	0.18	0.15	0.15	−0.02
Soil temperature (70 cm)	**0.56**	0.18	0.19	0.17	0.13	0.01
Soil moisture (5 cm)	**1**	**0.33**	**0.46**	**0.55**	**0.36**	0.02
Soil moisture (15 cm)		**1**	**0.90**	**0.76**	**0.63**	**0.86**
Soil moisture (30 cm)			**1**	**0.83**	**0.77**	**0.82**
Soil moisture (40 cm)				**1**	**0.65**	**0.60**
Soil moisture (50 cm)					**1**	**0.60**
Soil moisture (60 cm)						**1**

[a]The correlation method is the Pearson's method. Numbers in bold indicate significant correlation ($p < 0.05$). AQ means antecedent flow and is the total flow during the preceding 159, 7, and 30 days, as indicated.

dependent on soil temperature at 30 cm depth; that is, $f(t)$ takes the form $f(t) = \beta_{1,cf2} \sum_{i=t}^{t-159} Q_i + \beta_{0,cf2}$, and $c_d(t)$ takes the form $c_d(t) = \gamma_{1,\mathrm{cf2}} T_{\mathrm{soil}}(t) + \gamma_{0,\mathrm{cf2}}$.

[50] Other models are, of course, possible, but faced with so many possible model combinations, we made a judgment of models that we felt most likely to improve performance. Since we used linear functions in the models, we chose soil temperature measured in the upper soil, which correlated most strongly with the model residuals using Pearson's correlation. We did not use the soil temperature at 5, 10, or 15 cm since these levels are always above the groundwater table and are not hydrologically connected with the stream. Instead, we used modeled soil temperature at 30 cm.

3.3.1. Model Conditioning

[51] RIM_{static} displayed the same residual pattern in model conditioning for the period 2000–2006 as in the whole period evaluation with overestimation during spring flood and underestimation during summer and early fall. Maximum $E_{1,adj}$ was 0.22, while the maximum NS_{adj} was 0.42. The correlation between the f and the c_d parameters was high ($\rho = -0.99$, $p < 0.001$), and the model was sensitive to both parameters judging from the GSA.

[52] All dynamic versions of RIM performed better than RIM_{static} (Table 5 and Figure 4a), but there were differences among the different versions, with RIM_{temp} performing best overall on average. All dynamic models both underpredicted and overpredicted during spring flood. In the rest of the year, there was no clear pattern judging from model

Table 5. Performance of Different Versions of RIM in Model Conditioning Measured With the $E_{1,adj}$ and NS_{adj} Performance Measures, Maximum Absolute Error (MaxAE), Bias (BIAS), Median Absolute Error (MedAE), and Median Absolute Percent Error (PE)[a]

Model	$E1_{adj}$	NS_{adj}	MaxAE	BIAS	MedAE	PE
RIM_{static}	0.20 (0.01)	0.41 (0.01)	21.5 (1.0)	0.03 (0.12)	2.63 (0.07)	17.6 (0.7)
RIM_{temp}	0.42 (0.03)	0.69 (0.03)	11.4 (1.3)	0.03 (0.32)	1.88 (0.16)	12.9 (1.2)
RIM_{ant}	0.36 (0.03)	0.57 (0.03)	16.9 (1.9)	0.05 (0.27)	1.95 (0.17)	13.1 (1.0)
RIM_{f}	0.38 (0.05)	0.65 (0.04)	12.9 (2.2)	0.04 (0.39)	2.06 (0.28)	13.9 (1.8)
RIM_{cf1}	0.34 (0.05)	0.58 (0.06)	16.0 (2.8)	−0.01 (0.48)	2.12 (0.28)	14.4 (1.7)
RIM_{cf2}	0.38 (0.04)	0.64 (0.04)	14.6 (2.6)	−0.02 (0.33)	2.06 (0.27)	13.9 (1.6)

[a]All values are averages of the performance of the 200 behavioral parameter sets with standard deviation give in parentheses. The unit for MaxAE, MedAE, and BIAS is mg L^{-1}.

residuals, except for RIM_{ant}, which tended to underpredict during summer and fall. Both $E_{1,adj}$ and NS_{adj} for RIM_{temp} were higher than for RIM_{static}. The highest $E_{1,adj}$ was 0.46, while the maximum NS_{adj} was 0.73. The median absolute error (1.9 mg L^{-1}) and the percent error (13%) were substantially lower than for RIM_{static} (2.6 mg L^{-1} and 18%). All RIM_{temp} model parameters correlated with each other, but the highest correlation was between the c_d and β_0 parameters ($\rho = -0.99$, $p < 0.001$). In addition, the results of the GSA indicated that the model was sensitive to all parameters, especially to the β_0 parameter, i.e., the main determinant of the shape of the soil solution profile.

[53] RIM_{ant} also performed better than RIM_{static} but not as well as RIM_{temp}. The maximum $E_{1,adj}$ was 0.40, while the maximum NS_{adj} was 0.61. The c_d and β_0 parameters correlated strongly ($\rho = -0.98$, $p < 0.001$), while the other parameter combinations correlated weakly. RIM_{ant} was sensitive to all the parameters, with the β_0 parameter being the most sensitive and β_1 being the least sensitive.

[54] The maximum $E_{1,adj}$ for RIM_f was 0.47, while the maximum NS_{adj} was 0.73. The highest parameter correlations was between the c_d and the β_0 parameters ($\rho = -0.99$, $p < 0.001$) and between the β_1 and the β_2 parameters ($\rho = -0.76$, $p < 0.001$). Results from the GSA indicated that the model was sensitive to all the parameters, with the β_0 parameter being the most sensitive.

[55] The RIM_{cf1} and RIM_{cf2} models had similar dynamics in the conditioning period. The maximum $E_{1,adj}$ for RIM_{cf1} was 0.46, and the maximum NS_{adj} was 0.71. RIM_{cf2} had slightly higher maximum ($E_{1,adj} = 0.47$, $NS_{adj} = 0.72$) and mean values (Table 5). All parameters for both models correlated significantly, but the highest correlation was between the β_0 parameter and the γ_0 parameter for both models ($\rho = -0.98$, $p < 0.001$). According to the GSA, RIM_{cf1} was sensitive to all the parameters. RIM_{cf2} was insensitive to the γ_1 parameter but was sensitive to the other parameters. Both models turned out to be most sensitive to the β_0 parameter, i.e., the main determinant of the shape of the soil solution profile.

[56] None of the models performed better if the a and b parameters for hydrology were included in the model conditioning. The parameters of the different models were generally well constrained when the a and b parameters were not included in the conditioning, indicating limited equifinality (Figure 5).

3.3.2. Model Testing

[57] The results of testing the different RIM models for the period 1993–1999 were similar to the results in the conditioning period (Figure 4 and Table 6), with the exception of the RIM_{ant} model, which performed relatively better in the testing period. RIM_{temp} still had the highest average performance ($p < 0.05$), and RIM_{static} had the lowest

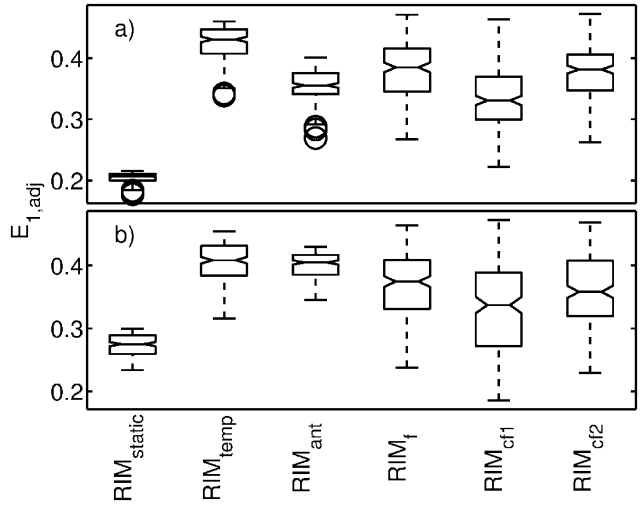

Figure 4. Model performance measured with the $E_{1,adj}$ performance measure in (a) the conditioning period (2000–2006) and (b) the test period (1993–1999).

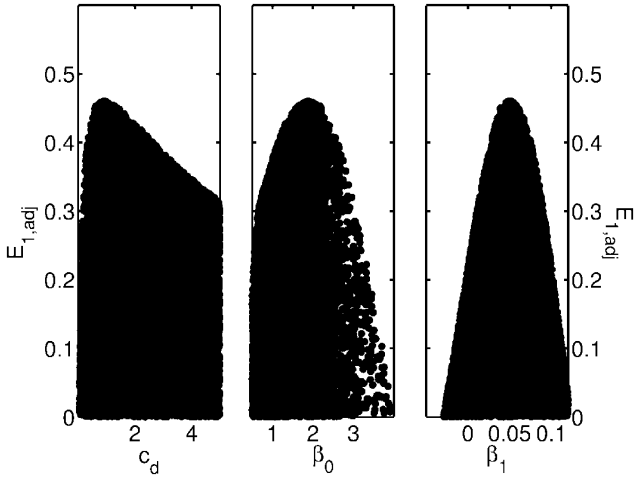

Figure 5. Model response to different parameter sets, so-called dotty plots in generalized likelihood uncertainty estimation (GLUE) terminology, for RIM_{temp} using performance measure $E_{1,adj}$.

Table 6. Performance of Different Versions of RIM in Model Testing Measured With the $E_{1,adj}$ and NS_{adj} Performance Measures, Maximum Absolute Error (MaxAE), Bias (BIAS), Median Absolute Error (MedAE), and Median Absolute Percent Error (PE)[a]

Model	$E1_{adj}$	NS_{adj}	MaxAE	BIAS	MedAE	PE
RIM_{static}	0.27 (0.02)	0.42 (0.01)	30.8 (0.6)	−0.59 (0.18)	2.99 (0.16)	23.6 (1.5)
RIM_{temp}	0.40 (0.03)	0.58 (0.03)	25.4 (1.2)	−0.06 (0.37)	2.14 (0.23)	17.4 (2.0)
RIM_{ant}	0.40 (0.02)	0.56 (0.02)	27.6 (1.2)	0.07 (0.34)	2.18 (0.23)	18.1 (1.6)
RIM_{f}	0.37 (0.05)	0.54 (0.05)	26.2 (1.6)	−0.16 (0.48)	2.30 (0.31)	18.6 (2.4)
RIM_{cf1}	0.33 (0.07)	0.51 (0.06)	27.7 (1.9)	−0.28 (0.58)	2.61 (0.51)	21.5 (4.0)
RIM_{cf2}	0.36 (0.05)	0.54 (0.05)	26.7 (1.8)	−0.22 (0.42)	2.41 (0.33)	19.2 (2.6)

[a]All values are averages of the performance of the 200 behavioral parameter sets with standard deviation given in parentheses. The unit for MaxAE, MedAE and BIAS is mg L^{-1}.

average performance measured with the $E_{1,adj}$, NS_{adj}, and MaxAE likelihood measures ($p < 0.05$) and according to the median absolute error and the percent error. RIM_{ant} had almost identical mean performance to RIM_{temp} on the basis of $E_{1,adj}$ (Figure 4b), but it had lower performance when measured with the other measures ($p < 0.05$). In addition, RIM_{temp} and RIM_{ant} had mean BIAS closest to zero.

[58] The differences in $E_{1,adj}$, NS_{adj}, and MaxAE between RIM_f and RIM_{cf2} were not significantly different, but the performance for both models was significantly lower than for RIM_{temp}. It was also lower than RIM_{ant} for $E_{1,adj}$, but the differences in NS_{adj} were not significant. RIM_{cf1} did not perform as well as the other models, except for RIM_{static}, measured with $E_{1,adj}$ or NS_{adj} ($p < 0.05$). The difference in BIAS between RIM_{temp} and RIM_{ant} or RIM_f was not significant. In addition, BIAS for RIM_f did not reveal a significant difference to RIM_{temp}, RIM_{cf1}, or RIM_{cf2}. RIM_{static} had the most negative BIAS on average.

[59] Although RIM_f, RIM_{cf1}, and RIM_{cf2} did not perform as well as RIM_{temp} on the basis of likelihood distribution statistics, they had higher maximum performance measured with $E_{1,adj}$ and NS_{adj} (Figure 4b). Time series plots of model predictions based on the RIM_{static} and RIM_{temp} models are shown in Figure 6.

[60] All dynamic versions of RIM simulated soil solution concentration profiles similar to observed data, as shown for RIM_{temp} in Figure 7, and had lower mean absolute error than RIM_{static} (Table 7). RIM_f and RIM_{temp} had the lowest mean absolute error between observed and simulated soil solution concentrations.

4. Discussion

[61] Seasonal variability of [DOC] in soils and streams is commonly observed [*Dawson et al.*, 2008; *Fröberg et al.*, 2006; *Köhler et al.*, 2008; *Lumsdon et al.*, 2005; *Seibert et al.*, 2009; *Ågren et al.*, 2008b]. While much of the short-term variability in [DOC] is dependent on discharge [*Hope et al.*, 1994; *Laudon et al.*, 2004a], it is clear that there are other important processes affecting the intra-annual DOC dynamics. Unraveling the causes of this short-term variability is crucial to accurate simulation of [DOC] variability, for example, under different climatic conditions.

[62] The results clearly show (Figure 6) that the effect of flow pathways on DOC dynamics alone, as conceptualized in RIM_{static}, is insufficient to explain the observed variability in DOC concentrations in the Västrabäcken stream. Although the model could simulate much of the dynamics, it missed the magnitude of the concentration increase of many of the large-flow peaks during summer and fall (Figure 3b). Many of the largest deviations occurred during high-flow events such as spring flood or summer and fall storms. These large residuals were consistent with the model assumption that soil solution [DOC] increases exponentially toward the soil surface. A deviation in the modeled soil solution profile from the "real" profile will have a relatively larger effect at high flow and high groundwater table position than at low flow. *Köhler et al.* [2009] also showed that a vertical shift of the flow paths in the riparian soil would have significant effects on the DOC dynamics in the stream.

[63] The strong relationship between the back-calculated f parameters, f_{back}, and the RIM_{static} residuals (Figure 3a) indicated that if one could describe the variability in f, it would be possible to reduce the residuals and improve the simulation of stream DOC. Hence, it would seem realistic to improve DOC modeling by varying f, i.e., the riparian soil solution concentration profile, if one could identify the underlying processes responsible for the variability in riparian soil solution [DOC]. The highest correlations between RIM_{static} residuals and our candidate drivers for [DOC] variability were obtained for soil temperature and antecedent flow. The stepwise regression also indicated that soil temperature and antecedent flow were significant predictors of the RIM_{static} residuals. Incorporating functions of these drivers in a dynamic RIM model structure improved model performance substantially (Table 6 and Figure 4).

[64] Antecedent flow could have an effect on DOC dynamics through flushing of DOC stored in the soil [*Boyer et al.*, 2000; *Hornberger et al.*, 1994]. However, one would then expect antecedent flow during a short term (for instance, 14 days) to be the most influential. On the contrary, we found that the cumulative flow during the preceding 5 months (159–168 days) had the largest effect on [DOC] variability, while the 14 day antecedent flow correlated weakly. This could indicate a long-term moisture effect on the production or mobilization of DOC [*Christ and David*, 1996; *Kalbitz et al.*, 2000]. The 159 day antecedent flow also correlated with observed soil moisture ($r = 0.66$–0.31, $p < 0.05$). It is, however, known that, e.g., primary productivity in these systems is generally not limited by water availability [*Bergh et al.*, 1999; *Phillips et al.*, 2001]. In addition, the antecedent flow covaried strongly with soil temperature as well (Table 4). The antecedent flow effect we found could thus have been an artifact caused by a seasonal signal in the discharge data that coincided with the soil temperature variability.

[65] Soil moisture has commonly been assumed to be an important driver for DOC variability [*Christ and David*,

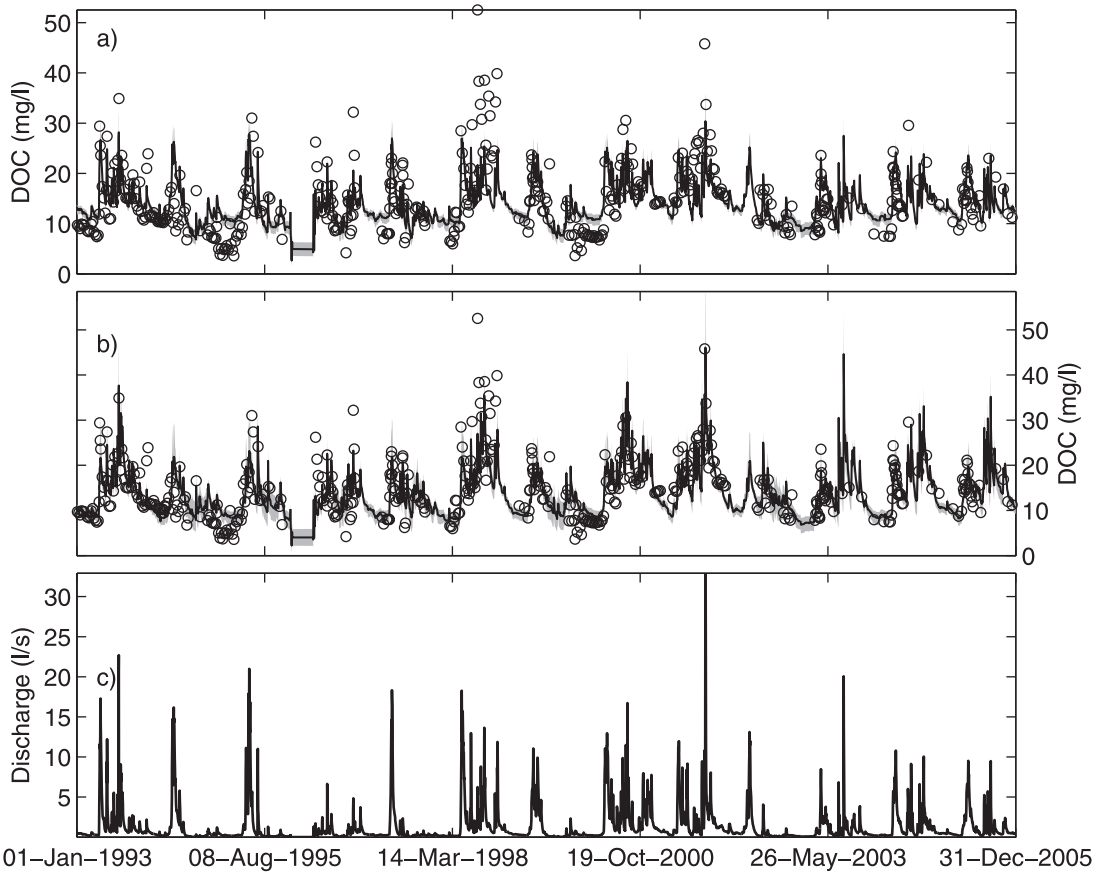

Figure 6. Time series plot of observed and simulated DOC concentrations in Västrabäcken using (a) RIM_{static} and (b) RIM_{temp}. (c) Time series plot of observed discharge in Västrabäcken. The solid line in Figures 6a and 6b is the median of all 200 top-performing parameter sets, the gray area is the uncertainty bounds using the same parameter sets, and circles are observed concentrations. The flat-looking area in the beginning of 1996 is caused by a lack of discharge data because of malfunctioning equipment.

1996], and it was also indicated by the stepwise regression analysis to be a significant predictor. However, it was not incorporated in any of the models since the correlation with model residuals was weak and often not significant. In addition, the distribution of soil moisture was bimodal, which artificially can give the impression of a strong correlation, although the factual relationship is weak, especially if using the Pearson's correlation. The persistence of the soil moisture could be even more important than the soil moisture as such [*Köhler et al.*, 2009].

[66] Recently, sulfate has been hypothesized to cause both short-term [*Clark et al.*, 2005] and long-term [*Evans et al.*, 2006; *Monteith et al.*, 2007] variability in DOC concentrations. We found a weak negative correlation between sulfate concentrations in the stream and model residuals, which could support the sulfate influence on [DOC] variability. The relationship was, however, weak and controlled by a few outliers, which makes further interpretation difficult. Additional exploration revealed that when soil temperature is accounted for, no significant effect of sulfate on [DOC] variability remained (data not shown). This could be a result of sulfate correlating with both discharge and soil temperature (Table 4).

[67] Although all the dynamic versions of RIM performed better than RIM_{static}, there were differences among the models. The results revealed that models where only the shape of the soil solution profile varied had, on average, higher performance than models where both the shape and the depth-averaged soil [DOC] varied (Figure 4). The highest average performance was achieved with RIM_{temp}, where variability in the modeled soil solution profile was only dependent on soil temperature. The variability in RIM_{temp}-simulated soil solution concentrations was also more similar than RIM_{static}-simulated concentrations to observed variability in riparian soil [DOC] (Figure 7). Since the soil solution data were not used in conditioning the models, this could be taken as evidence for the model simulating stream [DOC] for the right reasons. RIM_{ant} (in which the variability in soil solution profile was dependent on the 159 day antecedent flow) performed considerably better than the static model but not as well as RIM_{temp}. These results suggest that soil temperature is the main driver for the variability in soil solution [DOC]. Allowing the soil solution profile to vary on the basis of both soil temperature and antecedent flow (RIM_f) did not improve model results, which is expected given the correlation of soil temperature with the 159 day antecedent flow. Allowing the base concentration, i.e., the depth-averaged soil solution concentration, to vary (RIM_{cf1} and RIM_{cf2}) did not improve model results on average either. This could be a

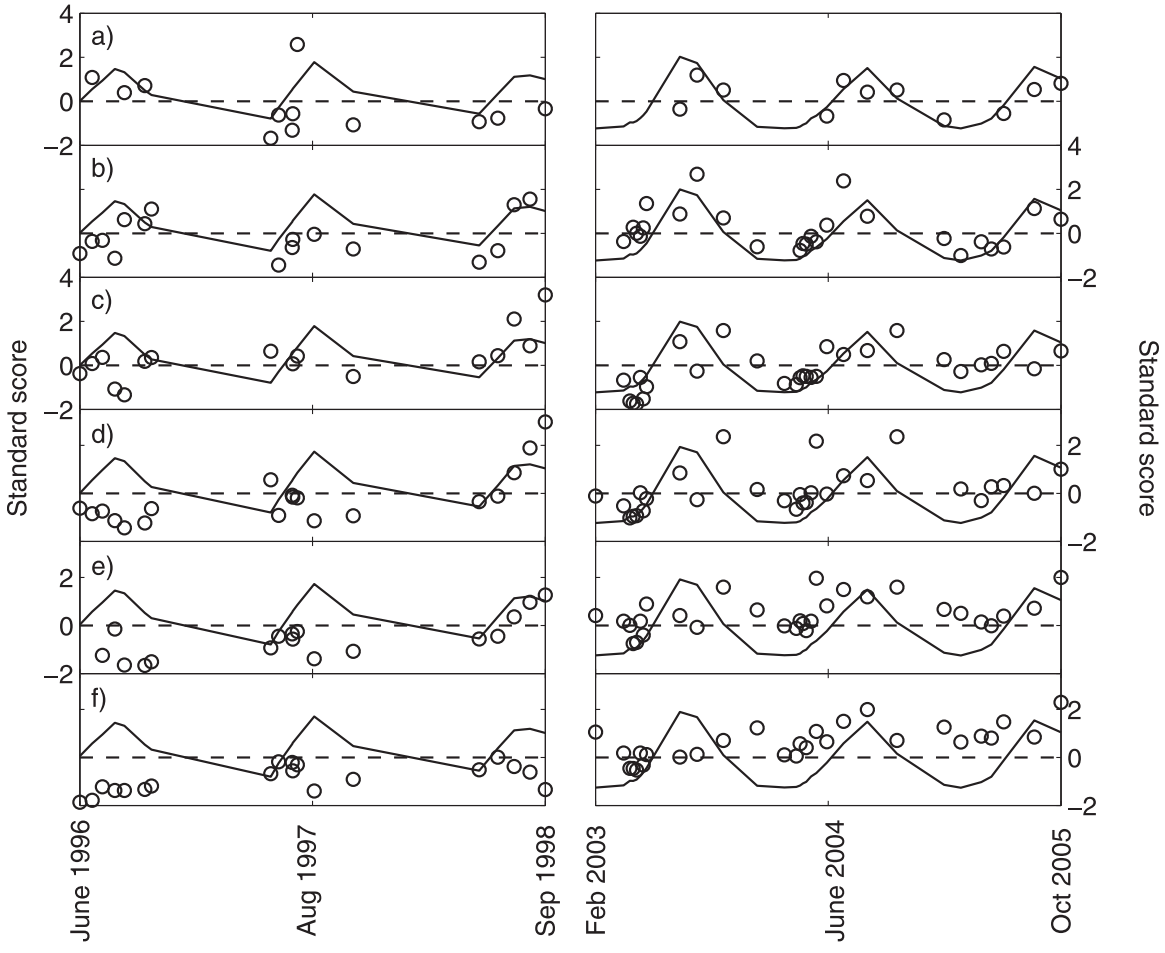

Figure 7. Standard scores of observed (circles), RIM_{static}-simulated (dashed line), and RIM_{temp}-simulated (solid line) DOC concentrations in riparian soil solution at (a) 10, (b) 25, (c) 35, (d) 45, (e) 55, and (f) 65 cm depth.

result of the observed covariation between soil temperature and antecedent flow, but the results also indicate that the riparian soil [DOC] variability is sufficiently captured by modeling the shape of the profile alone (Figure 7). That said, it is interesting to see that the performance differences among the dynamic models are small. Since RIM_{temp} has the highest performance on average and with just one parameter more than RIM_{static}, it is the most parsimonious hypothesis for the functioning of the system.

Table 7. Mean Absolute Error (MAE) Between Observed and Simulated Soil Solution Concentrations for the Different Versions of RIM[a]

Model	MAE
RIM_{static}	16.9
RIM_{temp}	13.4
RIM_{ant}	15.2
RIM_{f}	13.1
RIM_{cf1}	16.6
RIM_{cf2}	15.6

[a]All values are mean values for all sampling occasions at all depths. The unit for MAE is mg L^{-1}.

[68] RIM_f, RIM_{cf1}, and RIM_{cf2} had higher maximum values for $E_{1,adj}$ (Figure 4b), indicating that those models could perform better than RIM_{temp} if only the parameter space were covered in finer detail. Also, the variance in the likelihood measures was higher for the models with more parameters. This could be a result of the higher dimensionality of the parameter spaces of these models compared to models with fewer parameters. Adding an extra parameter, say from n to $n + 1$ parameters, would actually require an $i^{1/n}$ (i being the number of iterations for n parameters) increase in the number of iterations in the Monte Carlo simulation to cover the parameter space in equal detail. However, changing the number of iterations by an order of magnitude (from 20,000 to 200,000) increased the maximum NS_{adj} for RIM_f by about 1%, so we believe that the results in Figure 4 are a fair summary of the performance of the different models.

[69] There would be an advantage to use a model with fewer parameters if computer resources are scarce or costly besides the problem with overparameterization. The strong correlation between some of the parameters in the different models indicated that a more efficient sampling scheme than random sampling could be employed in the Monte Carlo simulation. One method could be to use a copula approach [*Beven*, 2008], which probably would be computationally

more efficient. It remains to be shown, however, if the parameter correlations are dependent on the time periods and catchment used in conditioning and testing of the models.

[70] Although the performance improved with the addition of a function of soil temperature to RIM, the models still did not perform well in some periods (Figure 6b). This could be due to large uncertainties in observed and modeled input data or evaluation data. It could also be a timing issue where the time resolution of input data (daily) is too coarse to fully correspond to the short-term temporal variability in [DOC] revealed by the "snapshots" in time represented by the grab samples. However, a more conceivable explanation would be shortcomings in the model structure. RIM is a simplified model ignoring many of the processes that are thought to be important for DOC dynamics. These include, for example, ionic strength and pH effects on organic carbon solubility [*Evans et al.*, 2008; *Monteith et al.*, 2007], sorption dynamics in the soil [*Jardine et al.*, 1989], microbial degradation [*Berggren et al.*, 2007], and topographic effects on the [DOC] variation in the stream. We also used the effective model parameter values and therefore assumed a homogeneous distribution of riparian concentrations of organic carbon and groundwater flow in the catchment. In doing this we ignored the spatial variability and any possible "hot spots" for organic carbon export [*Vidon et al.*, 2010]. Temporal changes in the dominant source area have previously been suggested as explanations for temporal [DOC] variability [*Laudon et al.*, 2004a; *Ågren et al.*, 2008a]. We used RIM as a lumped model, but it is possible to use RIM as a distributed model to account for spatial variability [*Grabs*, 2010]. Future development of the RIM concept would probably benefit from a representation of spatial heterogeneity in riparian soil solution concentrations and a representation of ionic strength effects on organic carbon dynamics. It is also of uttermost importance to test the RIM concept in more catchments with different landscape characteristics and climate.

[71] We see the major value of the RIM approach in efficiently accounting for flow effects with a physically meaningful conceptual structure to facilitate more refined analyses of the data. An excellent example of this is the analysis by *Ågren et al.* [2010], who calibrated RIM_{static} to a decade of spring floods to remove flow effects. They then used multivariate analysis on the residuals to identify how the weather conditions during the preceding winter and growing season influenced the interannual variability in spring flood DOC. This control on what happens during a few weeks of the year would be difficult to identify when working with overall performance for the entire year. However, we think that this study has shown that RIM is a useful tool to explore controls on stream water chemistry also at longer time scales.

[72] One should, of course, be careful about drawing too large conclusions about mechanistic causes from a modeling study. However, considering the different models as hypotheses about the system, one could speculate about the mechanisms behind the soil temperature effect on [DOC] in the soil and stream water. Soil temperature has been shown to affect the production as well as the degradation of DOC [*Christ and David*, 1996; *Kalbitz et al.*, 2000]. However, laboratory experiments have also indicated that there is no temperature effect on the extraction of DOC from soil samples [*Jones and Willett*, 2006]. This suggests that there may be some other processes, which are either sensitive to or just correlated in time with soil temperature variability, that are responsible for the seasonal signal in the [DOC] dynamics and not soil temperature as such. DOC production in soils has been hypothesized to be controlled by photosynthesis [*Giesler et al.*, 2007] as well as microbial degradation of organic matter [*Kalbitz et al.*, 2000], both of which are affected by temperature. In addition, the results in this study indicated that there is substantial variation in model performance in winter months (Figure 3c). This variability cannot be attributed to soil temperature variability because soil temperature is nearly constant during that period of the year. The largest residuals in winter months are probably during the spring flood, as indicated by the flow regimes (high and medium flows; see Figure 3). A possible cause of these deviations could be the severity of the winter, as shown by *Ågren et al.* [2010].

[73] Air temperature and runoff are expected to increase in many boreal areas in the future [*Andreasson et al.*, 2004; *Kjellström et al.*, 2011]. However, having this temperature rise be reflected in the soil temperature conditions is dependent on other factors such as snow cover extent. Increased future runoff has been hypothesized to lead to higher DOC concentrations in boreal surface waters [*Erlandsson et al.*, 2008], but should the future bring higher soil temperatures, the results from this study indicate that [DOC] could increase more than predicted from changes in discharge alone.

5. Conclusions

[74] Here we show how the riparian flow-concentration integration model (RIM), a simple dynamic model of riparian zone flows and concentrations, can be used to simulate first- and second-order controls on stream water DOC. The static RIM satisfactorily simulated the first-order hydrological controls on stream water DOC but was unable to reproduce the seasonality in the DOC-discharge relationship that is commonly observed in boreal systems. Adding a function of soil temperature in a dynamic RIM improved the performance substantially, providing good simulations with the first-order control of hydrology and second-order effect of soil temperature on stream water DOC. While discharge is the first-order control on organic carbon variability in many streams and rivers, the results from this study indicated that riparian soil temperature can serve as a proxy for the seasonality of DOC-flow patterns in this boreal stream. The RIM approach efficiently incorporates this second-order effect and lays the groundwork for further elucidation of the processes influencing the DOC dynamics. The approach used in this study, the use of RIM to control for flow-related changes in stream chemistry in order to reveal other factors, may be a simple method to analyze and model different constituents of stream water chemistry in different types of catchments.

[75] **Acknowledgments.** The authors would like to thank the staff at the Svartberget Research Station and the Krycklan Catchment Study for their help and support. Financial support for this work was provided by Formas and Mistra Future Forests. The long-term data set has been funded by Formas, the Swedish EPA, and the Swedish Research Council. The authors would also like to thank three anonymous reviewers and the associate editor for constructive comments on an early version of this paper.

References

Ågren, A., I. Buffam, M. Berggren, K. Bishop, M. Jansson, and H. Laudon (2008a), Dissolved organic carbon characteristics in boreal streams in a forest-wetland gradient during the transition between winter and summer, *J. Geophys. Res.*, *113*, G03031, doi:10.1029/2007JG000674.

Ågren, A., et al. (2008b), Seasonal and runoff-related changes in total organic carbon concentrations in the River Ore, northern Sweden, *Aquat. Sci.*, *70*, 21–29.

Ågren, A., et al. (2010), Regulation of stream water dissolved organic carbon (DOC) concentrations during snowmelt; the role of discharge, winter climate and memory effects, *Biogeosciences*, *7*, 2901–2913.

Andreasson, J., et al. (2004), Hydrological change—Climate change impact simulations for Sweden, *Ambio*, *33*, 228–234.

Berggren, M., H. Laudon, and M. Jansson (2007), Landscape regulation of bacterial growth efficiency in boreal freshwaters, *Global Biogeochem. Cycles*, *21*, GB4002, doi:10.1029/2006GB002844.

Bergh, J., et al. (1999), The effect of water and nutrient availability on the productivity of Norway spruce in northern and southern Sweden, *For. Ecol. Manage.*, *119*, 51–62.

Bertilsson, S., and L. J. Tranvik (2000), Photochemical transformation of dissolved organic matter in lakes, *Limnol. Oceanogr.*, *45*, 753–762.

Beven, K., (2008), *Environmental Modelling: An Uncertain Future?*, 304 pp., Routledge, London.

Beven, K. J., and A. Binley (1992), The future of distributed models: Model calibration and uncertainty prediction, *Hydrol. Processes*, *6*, 279–298.

Bishop, K., (1991), Episodic increases in stream acidity, catchment flow pathways and hydrograph separation, Ph.D. thesis, 246 pp., Univ. of Cambridge, Cambridge, U. K.

Bishop, K. H., et al. (1990), The origins of acid runoff in a hillslope during storm events, *J. Hydrol.*, *116*, 35–61.

Bishop, K., et al. (1994), Identification of the riparian sources of aquatic dissolved organic carbon, *Environ. Int.*, *20*, 11–19.

Bishop, K., et al. (1995), Terrestrial sources of methylmercury in surface waters: The importance of the riparian zone on the Svartberget catchment, *Water Air Soil Pollut.*, *80*, 435–444.

Bishop, K., et al. (2004), Resolving the double paradox of rapidly mobilized old water with highly variable responses in runoff chemistry, *Hydrol. Processes*, *18*, 185–189.

Boyer, E. W., et al. (2000), Effects of asynchronous snowmelt on flushing of dissolved organic carbon: A mixing model approach, *Hydrol. Processes*, *14*, 3291–3308.

Choi, H. T., and K. Beven (2007), Multi-period and multi-criteria model conditioning to reduce prediction uncertainty in an application of TOPMODEL within the GLUE framework, *J. Hydrol.*, *332*, 316–336.

Christ, M. J., and M. B. David (1996), Temperature and moisture effects on the production of dissolved organic carbon in a Spodosol, *Soil Biol. Biochem.*, *28*, 1191–1199.

Clark, J. M., et al. (2005), Influence of drought-induced acidification on the mobility of dissolved organic carbon in peat soils, *Global Change Biol.*, *11*, 791–809.

Clarke, R. T. (2008a), A critique of present procedures used to compare performance of rainfall-runoff models, *J. Hydrol.*, *352*, 379–387.

Clarke, R. T. (2008b), Issues of experimental design for comparing the performance of hydrologic models, *Water Resour. Res.*, *44*, W01409, doi:10.1029/2007WR005927.

Currie, W. S., and J. D. Aber (1997), Modeling leaching as a decomposition process in humid montane forests, *Ecology*, *78*(6), 1844–1860.

Dawson, J., et al. (2008), Influence of hydrology and seasonality on DOC exports from three contrasting upland catchments, *Biogeochemistry*, *90*, 93–113.

Dosskey, M. G., and P. M. Bertsch (1994), Forest sources and pathways of organic matter transport to a blackwater stream: A hydrologic approach, *Biogeochemistry*, *24*, 1–19.

Eimers, M. C., et al. (2008), Influence of seasonal changes in runoff and extreme events on dissolved organic carbon trends in wetland- and upland-draining streams, *Can. J. Fish. Aquat. Sci.*, *65*, 796–808.

Erlandsson, M., et al. (2008), Thirty-five years of synchrony in the organic matter concentrations of Swedish rivers explained by variation in flow and sulphate, *Global Change Biol.*, *14*, 1191–1198.

Eshleman, K. N., and H. F. Hemond (1985), The role of organic-acids in the acid-base status of surface waters at Bickford watershed, Massachusetts, *Water Resour. Res.*, *21*, 1503–1510, doi:10.1029/WR021i010p01503.

Evans, C. D., et al. (2006), Alternative explanations for rising dissolved organic carbon export from organic soils, *Global Change Biol.*, *12*, 2044–2053.

Evans, C., et al. (2008), Does elevated nitrogen deposition or ecosystem recovery from acidification drive increased dissolved organic carbon loss from upland soil? A review of evidence from field nitrogen addition experiments, *Biogeochemistry*, *91*, 13–35.

Fiebig, D. M., et al. (1990), Soil water in the riparian zone as a source of carbon for a headwater stream, *J. Hydrol.*, *116*, 217–237.

Findlay, S. E. G. (2005), Increased carbon transport in the Hudson River: Unexpected consequence of nitrogen deposition?, *Frontiers Ecol. Environ.*, *3*, 133–137.

Finlay, J., J. Neff, S. Zimov, A. Davydova, and S. Davydov (2006), Snowmelt dominance of dissolved organic carbon in high-latitude watersheds: Implications for characterization and flux of river DOC, *Geophys. Res. Lett.*, *33*, L10401, doi:10.1029/2006GL025754.

Freeman, C., et al. (2001), Export of organic carbon from peat soils, *Nature*, *412*, 785–785.

Freeman, C., et al. (2004), Export of dissolved organic carbon from peatlands under elevated carbon dioxide levels, *Nature*, *430*, 195–198.

Fröberg, M., et al. (2006), Concentration and fluxes of dissolved organic carbon (DOC) in three Norway spruce stands along a climatic gradient in Sweden, *Biogeochemistry*, *77*, 1–23.

Futter, M. N., D. Butterfield, B. J. Cosby, P. J. Dillon, A. J. Wade, and P. G. Whitehead (2007), Modeling the mechanisms that control in-stream dissolved organic carbon dynamics in upland and forested catchments, *Water Resour. Res.*, *43*, W02424, doi:10.1029/2006WR004960.

Gadmar, T. C., et al. (2002), The merits of the high-temperature combustion method for determining the amount of natural organic carbon in surface freshwater samples, *Int. J. Environ. Anal. Chem.*, *82*, 451–461.

Giesler, R., et al. (2007), Production of dissolved organic carbon and low-molecular weight organic acids in soil solution driven by recent tree photosynthate, *Biogeochemistry*, *84*, 1–12.

Grabs, T., (2010), Water quality modeling based on landscape analysis: Importance of riparian hydrology, Ph.D. thesis, 40 pp, Stockholm Univ., Stockholm.

Grieve, I. C. (1991), A model of dissolved organic carbon concentrations in soil and stream waters, *Hydrol. Processes*, *5*, 301–307.

Gupta, H. V., S. Sorooshian, and P. O. Yapo (1998), Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, *Water Resour. Res.*, *34*, 751–763, doi:10.1029/97WR03495.

Haei, M., M. G. Öquist, I. Buffam, A. Ågren, P. Blomkvist, K. Bishop, M. Ottosson Löfvenius, and H. Laudon (2010), Cold winter soils enhance dissolved organic carbon concentrations in soil and stream water, *Geophys. Res. Lett.*, *37*, L08501, doi:10.1029/2010GL042821.

Hagedorn, F., et al. (2001), Contrasting dynamics of dissolved inorganic and organic nitrogen in soil and surface waters of forested catchments with Gleysols, *Geoderma*, *100*, 173–192.

Hinton, M. J., et al. (1998), Sources and flowpaths of dissolved organic carbon during storms in two forested watersheds of the Precambrian Shield, *Biogeochemistry*, *41*, 175–197.

Hope, D., et al. (1994), A review of the export of carbon in river water: Fluxes and processes, *Environ. Pollut.*, *84*, 301–324.

Hornberger, G. M., et al. (1994), Hydrological controls on dissolved organic carbon during snowmelt in the Snake River near Montezuma, Colorado, *Biogeochemistry*, *25*, 147–165.

Hruska, J., et al. (2001), Acid/base character of organic acids in a boreal stream during snowmelt, *Water Resour. Res.*, *37*, 1043–1056, doi:10.1029/2000WR900290.

Jansson, M., et al. (2007), Terrestrial carbon and intraspecific size-variation shape lake ecosystems, *Trends Ecol. Evol.*, *22*, 316–322.

Jardine, P. M., et al. (1989), Mechanisms of dissolved organic-carbon adsorption on soil, *Soil Sci. Soc. Am. J.*, *53*, 1378–1385.

Jones, D. L., and V. B. Willett (2006), Experimental evaluation of methods to quantify dissolved organic nitrogen (DON) and dissolved organic carbon (DOC) in soil, *Soil Biol. Biochem.*, *38*, 991–999.

Kalbitz, K., et al. (2000), Controls on the dynamics of dissolved organic matter in soils: A review, *Soil Sci.*, *165*, 277–304.

Karlsson, J., et al. (2009), Light limitation of nutrient-poor lake ecosystems, *Nature*, *460*, 506–509.

Kendall, K. A., et al. (1999), A hydrometric and geochemical approach to test the transmissivity feedback hypothesis during snowmelt, *J. Hydrol.*, *219*, 188–205.

Kirchner, J. W. (2003), A double paradox in catchment hydrology and geochemistry, *Hydrol. Processes*, *17*, 871–874.

Kjellström, E., et al. (2011), 21st century changes in the European climate: Uncertainties derived from an ensemble of regional climate model simulations, *Tellus, Ser. A*, *63*, 24–40.

Klemes, V., (1986), Operational testing of hydrological simulation models, *Hydrol. Sci. J.*, *31*, 13–24.

Kruskal, W. H., and W. A. Wallis (1952), Use of ranks in one-criterion variance analysis, *J. Am. Stat. Assoc.*, *47*, 583–621.

Köhler, S. J., I. Buffam, H. Laudon, and K. H. Bishop (2008), Climate's control of intra-annual and interannual variability of total organic carbon concentration and flux in two contrasting boreal landscape elements, *J. Geophys. Res.*, *113*, G03012, doi:10.1029/2007JG000629.

Köhler, S. J., et al. (2009), Dynamics of stream water TOC concentrations in a boreal headwater catchment: Controlling factors and implications for climate scenarios, *J. Hydrol.*, *373*, 44–56.

Lajtha, K., et al. (2005), Detrital controls on soil solution N and dissolved organic matter in soils: A field experiment, *Biogeochemistry*, *76*, 261–281.

Latch, D. E., and K. McNeill (2006), Microheterogeneity of singlet oxygen distributions in irradiated humic acid solutions, *Science*, *311*, 1743–1747.

Laudon, H., H. F. Hemond, R. Krouse, and K. H. Bishop (2002), Oxygen 18 fractionation during snowmelt: Implications for spring flood hydrograph separation, *Water Resour. Res.*, *38*(11), 1258, doi:10.1029/2002WR001510.

Laudon, H., et al. (2004a), Seasonal TOC export from seven boreal catchments in northern Sweden, *Aquatic Sci.*, *66*, 223–230.

Laudon, H., J. Seibert, S. Köhler, and K. Bishop (2004b), Hydrological flow paths during snowmelt: Congruence between hydrometric measurements and oxygen 18 in meltwater, soil water, and runoff, *Water Resour. Res.*, *40*, W03102, doi:10.1029/2003WR002455.

Laudon, H., et al. (2007), The role of catchment scale and landscape characteristics for runoff generation of boreal streams, *J. Hydrol.*, *344*, 198–209.

Legates, D. R., and G. J. McCabe Jr. (1999), Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation, *Water Resour. Res.*, *35*, 233–241, doi:10.1029/1998WR900018.

Löfvenius, M. O., et al. (2003), Snow and soil frost depth in two types of shelterwood and a clear-cut area, *Scand. J. For. Res.*, *18*, 54–63.

Lumsdon, D. G., et al. (2005), Model assessment of biogeochemical controls on dissolved organic carbon partitioning in an acid organic soil, *Environ. Sci. Technol.*, *39*, 8057–8063.

Michalzik, B., et al. (2003), Modelling the production and transport of dissolved organic carbon in forest soils, *Biogeochemistry*, *66*(3), 241–264.

Monteith, D. T., et al. (2007), Dissolved organic carbon trends resulting from changes in atmospheric deposition chemistry, *Nature*, *450*, 537–540.

Nash, J. E., and J. V. Sutcliffe (1970), River flow forecasting through conceptual models part I—A discussion of principles, *J. Hydrol.*, *10*, 282–290.

Neff, J. C., and G. P. Asner (2001), Dissolved Organic Carbon in Terrestrial Ecosystems: Synthesis and a Model, *Ecosystems*, *4*(1), 29–48.

Nyberg, L., et al. (2001), Soil frost effects on soil water and runoff dynamics along a boreal forest transect: 1. Field investigations, *Hydrol. Processes*, *15*, 909–926.

Phillips, N., et al. (2001), Effects of nutrition and soil water availability on water use in a Norway spruce stand, *Tree Physiol.*, *21*, 851–860.

Press, W. H., et al. (2007), *Numerical Recipes: The Art of Scientific Computing*, 3rd ed., 1235 pp., Cambridge Univ. Press, New York.

Rankinen, K., et al. (2004), A simple model for predicting soil temperature in snow-covered and seasonally frozen soil: Model description and testing, *Hydrol. Earth Syst. Sci.*, *8*, 706–716.

Ravichandran, M., (2004), Interactions between mercury and dissolved organic matter—A review, *Chemosphere*, *55*, 319–331.

Rodhe, A., (1987), The origin of streamwater traced by oxygen-18, Ph.D. thesis, Uppsala Univ., Uppsala, Sweden.

Rodhe, A., (1989), On the generation of stream runoff in till soils, *Nord. Hydrol.*, *20*, I–8.

Roulet, N., and T. R. Moore (2006), Environmental chemistry: Browning the waters, *Nature*, *444*, 283–284.

Scheffe, H., (1953), A method for judging all contrasts in the analysis of variance, *Biometrika*, *40*, 87–104.

Schulten, H. R., and M. Schnitzer (1993), A state of the art structural concept for humic substances, *Naturwissenschaften*, *80*, 29–30.

Seibert, J., and J. J. McDonnell (2002), On the dialog between experimentalist and modeler in catchment hydrology: Use of soft data for multicriteria model calibration, *Water Resour. Res.*, *38*(11), 1241, doi:10.1029/2001WR000978.

Seibert, J., K. Bishop, A. Rodhe, and J. J. McDonnell (2003), Groundwater dynamics along a hillslope: A test of the steady state hypothesis, *Water Resour. Res.*, *39*(1), 1014, doi:10.1029/2002WR001404.

Seibert, J., et al. (2009), Linking soil- and stream-water chemistry based on a riparian flow-concentration integration model, *Hydrol. Earth Syst. Sci.*, *13*, 2287–2297.

Spear, R. C., and G. M. Hornberger (1980), Eutrophication in Peel Inlet—II. Identification of critical uncertainties via generalized sensitivity analysis, *Water Res.*, *14*, 43–49.

Tipping, E., (2002), *Cation Binding by Humic Substances*, 434 pp., Cambridge Univ. Press, Cambridge, U. K.

Weiler, M., and J. J. McDonnell (2006), Testing nutrient flushing hypotheses at the hillslope scale: A virtual experiment approach, *J. Hydrol.*, *319*, 339–356.

Vidon, P., et al. (2010), Hot spots and hot moments in riparian zones: Potential for improved water quality management, *J. Am. Water Resour. Assoc.*, *46*, 278–298.

Worrall, F., et al. (2004), Trends in dissolved organic carbon in UK rivers and lakes, *Biogeochemistry*, *70*, 369–402.

Yang, W. C., et al. (2006), Effects of dissolved organic matter on permethrin bioavailability to *Daphnia* species, *J. Agric. Food Chem.*, *54*, 3967–3972.

Yurova, A., et al. (2008), Modeling the dissolved organic carbon output from a boreal mire using the convection-dispersion equation: Importance of representing sorption, *Water Resour. Res.*, *44*, W07411, doi:10.1029/2007WR006523.

K. Bishop, M. Futter, S. Köhler, and M. Winterdahl, Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, PO Box 7050, SE-750 07 Uppsala, Sweden. (mattias.winterdahl@vatten.slu.se)

H. Laudon, Department of Forest Ecology and Management, Swedish University of Agricultural Sciences, SE-901 83 Umeå, Sweden.

J. Seibert, Department of Geography, University of Zurich, CH-8057 Zurich, Switzerland.